

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2004, 06:53:17 ; Search time 38.7949 Seconds
(without alignments)
456.424 Million cell updates/sec

Title: US-09-803-918A-2
Perfect score: 1362
Sequence: 1 MKAADVTLAVLFTGSQARH.....SPKVSFLSALEBYTKKLNTQ 267

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/prodata/1/iaa/5A COMB.pep.*
- 2: /cgn2_6/prodata/1/iaa/5B COMB.pep.*
- 3: /cgn2_6/prodata/1/iaa/6A COMB.pep.*
- 4: /cgn2_6/prodata/1/iaa/6B COMB.pep.*
- 5: /cgn2_6/prodata/1/iaa/PCTUS COMB.pep.*
- 6: /cgn2_6/prodata/1/iaa/backfilesl.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1362	100.0	267	1 US-07-959-946-3	Sequence 3, Appli
2	1362	100.0	267	1 US-08-333-577-3	Sequence 3, Appli
3	1362	100.0	267	1 US-08-952-796-2	Sequence 2, Appli
4	1362	100.0	267	4 US-08-940-136-260	Sequence 260, App
5	1362	100.0	267	5 PCT-US92-08634-3	Sequence 3, Appli
6	1346	98.8	267	4 US-09-919-039-27	Sequence 27, Appl
7	1241	91.1	243	4 US-09-079-030-119	Sequence 119, App
8	1240.5	91.1	264	1 US-08-448-606-6	Sequence 6, Appli
9	1006	73.9	200	3 US-08-952-796-15	Sequence 15, Appl
10	328	24.1	64	2 US-08-292-870-1	Sequence 1, Appli
11	278.5	20.4	396	4 US-09-800-729-207	Sequence 207, App
12	269	19.8	382	4 US-09-800-729-206	Sequence 206, App
13	251	18.4	391	4 US-09-800-729-208	Sequence 208, App
14	229.5	16.9	397	4 US-09-079-030-123	Sequence 123, App
15	227	16.7	44	2 US-08-292-870-2	Sequence 2, Appli
16	200	14.7	366	4 US-09-750-580-3	Sequence 3, Appli
17	200	14.7	366	4 US-09-599-360B-89	Sequence 89, Appl
18	198	14.5	363	4 US-09-800-729-111	Sequence 111, App
19	198	14.5	366	4 US-09-800-729-212	Sequence 212, App
20	190	14.0	317	3 US-08-949-155-6	Sequence 6, Appli
21	190	14.0	317	3 US-09-819-964-6	Sequence 6, Appli
22	174.5	12.8	316	4 US-09-919-039-247	Sequence 247, App
23	171.5	12.6	317	1 US-07-709-949-2	Sequence 2, Appli
24	163	12.0	32	2 US-08-292-870-3	Sequence 3, Appli
25	147.5	10.8	239	4 US-09-108-006C-62	Sequence 62, Appl
26	145	10.6	242	4 US-09-543-681A-5933	Sequence 5933, Ap
27	133.5	9.8	1972	4 US-08-875-435B-3	Sequence 3, Appli

Query Match 100.0%; Score 1362; DB 1; Length 267;
Best Local Similarity 100.0%; Pred. No. 7.1e-109;

28	131	9.6	579	4	US-09-538-092-1284	Sequence 1284, Ap
29	131	9.6	585	4	US-09-462-951B-1	Sequence 1, Appli
30	131	9.6	1530	4	US-09-976-594-736	Sequence 736, App
31	130	9.5	1201	4	US-09-252-991A-32259	Sequence 32259, A
32	129.5	9.5	977	4	US-09-010-147B-18	Sequence 18, Appl
33	129	9.5	900	2	US-08-630-822A-62	Sequence 62, Appl
34	129	9.5	900	2	US-09-005-069-62	Sequence 62, Appl
35	129	9.5	900	3	US-09-171-156A-21	Sequence 21, Appl
36	129	9.5	900	4	US-09-004-730A-21	Sequence 21, Appl
37	129	9.5	900	4	US-08-981-799A-21	Sequence 21, Appl
38	128.5	9.4	1976	4	US-09-538-092-1078	Sequence 1078, Ap
39	127.5	9.4	691	4	US-10-140-002-16	Sequence 16, Appl
40	126	9.3	220	2	US-08-726-306A-29	Sequence 29, Appl
41	126	9.3	718	4	US-09-540-236-2753	Sequence 2753, Ap
42	126	9.3	1935	4	US-09-538-092-916	Sequence 916, App
43	125.5	9.2	1078	4	US-09-248-796A-20284	Sequence 20284, A
44	124	9.1	25	2	US-08-292-870-4	Sequence 4, Appli
45	123.5	9.1	683	6	5210183-3	Patent No. 5210183

ALIGNMENTS

RESULT 1
US-07-959-946-3
; Sequence 3, Application US/07959946
; Patent No. 5408038
; GENERAL INFORMATION:
; APPLICANT: Smith, Richard K.
; APPLICANT: Koduri, Raju
; APPLICANT: Young, Stephen G.
; APPLICANT: Witztum, Joseph L.
; APPLICANT: Curtiss, Linda K.
; TITLE OF INVENTION: Lipoprotein Assays Using Antibodies to a
; TITLE OF INVENTION: Pan Native Epitope and Recombinant Antigens
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dressler, Goldsmith, Shore, Sutker &
; ADDRESSEE: Milnamow, Ltd.
; STREET: 180 No. 5408038th Stetson, Suite 4700
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60601

COMPUTER TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/959,946
FILING DATE: 19921008
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/901,706
FILING DATE: 18-JUN-1992
ATTORNEY/AGENT INFORMATION:
NAME: Gamson, Edward P.
REGISTRATION NUMBER: 29,381
REFERENCE/DOCKET NUMBER:
TELEPHONE: (312)616-5400
TELEFAX: (312)616-5460
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 267 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-959-946-3

[illegible]

```

RESULT 2
US-08-333-577-3
; Sequence 3, Application US/08333577
; Patent No. 5786206
; GENERAL INFORMATION:
; APPLICANT: Smith, Richard K.
; APPLICANT: Koduri, Raju
; APPLICANT: Young, Stephen G.
; APPLICANT: Witztum, Joseph L.
; APPLICANT: Curtiss, Linda K.
; TITLE OF INVENTION: Lipoprotein Assays Using Antibodies to a
; Pan Native Epitope and Recombinant Antigens
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dressler, Goldsmith, Shore, Sutker &
; ADDRESSEE: Milnamow, Ltd.
; STREET: 180 No. 5786206th Stetson, Suite 4700
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/333,577
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Gamson, Edward P.
; REGISTRATION NUMBER: 29,381
; REFERENCE/DOCKET NUMBER: SCRF 234.0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312)616-5400
; TELEFAX: (312)616-5460
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 267 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-333-577-3

Query Match 100.0%; Score 1362; DB 1; Length 267;
Best Local Similarity 100.0%; Pred.No. 7.1e-109;
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKAAVLTAVLFLFGTSQARHFQWQDEPPQSPDWRVKOLATVYVDVLKDSGRDYVSQFEGS 60

```

Query Match 100.0%; Score 1362; DB 3; Length 267;
Best Local Similarity 100.0%; Pred. No. 7.1e-109;
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKAAVLTAVLFLTGSQARHFHQDEPPQSPWDRVKDLATVYVDLKDGRDYVSQFEGS 60
DB 1 MKAAVLTAVLFLTGSQARHFHQDEPPQSPWDRVKDLATVYVDLKDGRDYVSQFEGS 60

QY 61 ALGKQLNLKLLDNWDSVTSTFSKLRQLGPGVTQEFWMDNLEKETEGLRQEMSKDLEEVKAK 120
DB 61 ALGKQLNLKLLDNWDSVTSTFSKLRQLGPGVTQEFWMDNLEKETEGLRQEMSKDLEEVKAK 120

QY 121 VQPYLDDFOKKWOEMELVYRKVEPLRAELQEGARQKLHELOEKLSPGLGEEMDRARAHV 180
DB 121 VQPYLDDFOKKWOEMELVYRKVEPLRAELQEGARQKLHELOEKLSPGLGEEMDRARAHV 180

QY 181 DALRTHLAPYSDLRQRLAARLEALKENGARLAELYHAKATEHLSTLSEKAKPALEDLRQ 240
DB 181 DALRTHLAPYSDLRQRLAARLEALKENGARLAELYHAKATEHLSTLSEKAKPALEDLRQ 240

QY 241 GLLPVLESFKVSFLSALEYTKKLNQ 267
DB 241 GLLPVLESFKVSFLSALEYTKKLNQ 267

RESULT 4
US-08-940-136-260
; Sequence 260, Application US/08940136
; Patent No. 6518412
; GENERAL INFORMATION:
; APPLICANT: Daseux, Jean-Louis
; APPLICANT: Sekul, Renate
; APPLICANT: Buttner, Klaus
; APPLICANT: Cornut, Isabelle
; APPLICANT: Metz, Gunther
; APPLICANT: Dufourcq, Jean
; TITLE OF INVENTION: GENE THERAPY APPROACHES TO
; TITLE OF INVENTION: SUPPLY APOIPOPROTEIN A-I AGONISTS AND THEIR
; TITLE OF INVENTION: USE TO TREAT DYSLIPIDEMIC DISORDERS.
; NUMBER OF SEQUENCES: 274
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/940,136
; FILING DATE: 29-SEP-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 009196-0007-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-493-4935
; TELEFAX: 650-493-5556
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 260:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 267 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
US-08-940-136-260

Query Match 100.0%; Score 1362; DB 4; Length 267;
Best Local Similarity 100.0%; Pred. No. 7.1e-109;
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKAAVLTAVLFLTGSQARHFHQDEPPQSPWDRVKDLATVYVDLKDGRDYVSQFEGS 60
DB 1 MKAAVLTAVLFLTGSQARHFHQDEPPQSPWDRVKDLATVYVDLKDGRDYVSQFEGS 60

QY 61 ALGKQLNLKLLDNWDSVTSTFSKLRQLGPGVTQEFWMDNLEKETEGLRQEMSKDLEEVKAK 120
DB 61 ALGKQLNLKLLDNWDSVTSTFSKLRQLGPGVTQEFWMDNLEKETEGLRQEMSKDLEEVKAK 120

QY 121 VQPYLDDFOKKWOEMELVYRKVEPLRAELQEGARQKLHELOEKLSPGLGEEMDRARAHV 180
DB 121 VQPYLDDFOKKWOEMELVYRKVEPLRAELQEGARQKLHELOEKLSPGLGEEMDRARAHV 180

QY 181 DALRTHLAPYSDLRQRLAARLEALKENGARLAELYHAKATEHLSTLSEKAKPALEDLRQ 240
DB 181 DALRTHLAPYSDLRQRLAARLEALKENGARLAELYHAKATEHLSTLSEKAKPALEDLRQ 240

QY 241 GLLPVLESFKVSFLSALEYTKKLNQ 267
DB 241 GLLPVLESFKVSFLSALEYTKKLNQ 267

RESULT 5
PCT-US92-08634-3
; Sequence 3, Application PC/TUS9208634
; GENERAL INFORMATION:
; APPLICANT: Smith, Richard K.
; APPLICANT: Koduri, Raju
; APPLICANT: Young, Stephen G.
; APPLICANT: Witzum, Joseph L.
; APPLICANT: Curtiss, Linda K.
; TITLE OF INVENTION: Lipoprotein Assays Using Antibodies to a
; TITLE OF INVENTION: Pan Native Epitope and Recombinant Antigens
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Dressler, Goldsmith, Shore, Sutker &
; STREET: 180 North Stetson, Suite 4700
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/08634
; FILING DATE: 19921009
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/901,706
; FILING DATE: 18-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Gamson, Edward P.
; REGISTRATION NUMBER: 29,381
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312)616-5400
; TELEFAX: (312)616-5460
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 267 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear

; MOLECULE TYPE: protein
PCT-US92-08634-3

```
Query Match      100.0%; Score 1362; DB 5; Length 267;
Best Local Similarity 100.0%; Pred. No. 7.1e-109;
Matches 267; Conservative 0; Mismatches 0; Indels 0
```

1	MKAAVLTLAVFLTGSQA	BHFQODEPP	OS	PWDRVKDLATVV	DVILKDSGRD	VYVSQFEGS	60
1	MKAAVLTLAVFLTGSQA	BHFQODEPP	OS	PWDRVKDLATVV	DVILKDSGRD	VYVSQFEGS	60
61	ALGKQLNLKLLDNWDS	VTSTF	SKLREQLGPV	TQEF	FDNLLEKETEG	LRSQMSKDLEEV	KAK 120
61	ALGKQLNLKLLDNWDS	VTSTF	SKLREQLGPV	TQEF	FDNLLEKETEG	LRSQMSKDLEEV	KAK 120
121	VOPYLDDFOKKQW	EEMLYRQ	KVEPLR	ABELQEGARQ	KLHELOEKL	SPLGEEMRDRAR	AHV 180
121	VOPYLDDFOKKQW	EEMLYRQ	KVEPLR	ABELQEGARQ	KLHELOEKL	SPLGEEMRDRAR	AHV 180
181	DALRTHLAPYSDEL	RQRLAAR	LEALKENG	GARLAEYHAKA	THSLSTL	SEKAPPALEDLR	Q 240
181	DALRTHLAPYSDEL	RQRLAAR	LEALKENG	GARLAEYHAKA	THSLSTL	SEKAPPALEDLR	Q 240
241	GLLPVLESGFKVS	FLSAL	EEYTKKL	NTQ 267			
241	GLLPVLESGFKVS	FLSAL	EEYTKKL	NTQ 267			

RESULT 6

```

US-09-919-039-27
; Sequence 27, Application US/09919039
; Patent No. 6727066
; GENERAL INFORMATION:
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
; FILE REFERENCE: PA-0035 US
; CURRENT APPLICATION NUMBER: US/09/919,039
; CURRENT FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 60/222,113
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 401
; SOFTWARE: PERL Program
; SEQ ID NO 27
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6727066 2516070CD1
US-09-919-039-27

```

Query Match	98.8%;	Score 1346;	DB 4;	Length 267;
Best Local Similarity	98.9%;	Pred. NO. 1.7e-107;		
Matches 264;	Conservative 1;	Mismatches 2;	Indels 0;	Gaps 0
Qy	1	MKAAVLTAVLFLTGSQARHFWQDEPQSPDWRVKDLATVYVDVLKDSGRDVVSQFEGS	60	
Db	1	MKAAVLTAVLFLTGSQARHFWQDEPQSPDWRVKDLATVYVDVLKDSGRDVVSQFEGS	60	
Qy	61	ALGKQLNLKLLDNWDSVTSTFSKLEQLGPGVPTQEFWNLLEKETEGLEQMSKDLSEVKA	120	
Db	61	ALGKQLNLKLLDNWDSVTSTFSKLEQLGPGVPTQEFWNLLEKETEGLEQMSKDLSEVKA	120	
Qy	121	VQPYLDDFOKKQWQEMELYRQKVEPLRAELQEGARQKHLHELQKLSPLGSEMDRRAHV	180	
Db	121	VQPYLDDFOKKQWQEMELYRQKVEPLRAELQEGARQKHLHELQKLSPLGSEMDRRAHV	180	
Qy	181	DALRTHLAPYSDELQRRLAARLEALKENGSGARLAAYTHAKATEHLSTLSEKAKPALSDLRQ	240	
Db	181	ARVRTHLAPYSDELQRRLAARLEALKENGSGARLAAYTHAKATEHLSTLSEKAKPALSDLRQ	240	
Qy	241	GLLPVLESFKV3FLSALEYTKKLTNQ	267	

241 GLLPVLESFKVSFLSALEEYTKKLNQ 267

RESULT 7

```

US-09-079-030-119
; Sequence 119, Application US/09079030
; Patent No. 6635623
;
; GENERAL INFORMATION:
;
; APPLICANT: Guevera, Jr., Juan G.
; APPLICANT: Hoogerveen, Ron C.
; APPLICANT: Moore, Paul J.
; TITLE OF INVENTION: LIPOPROTEINS AS NUCLEIC ACID DELIVERY
; TITLE OF INVENTION: VECTORS FOR TRANSCRIPTION OF EUKARYOTIC CELLS
; NUMBER OF SEQUENCES: 229
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/079,030
; FILING DATE: Concurrently Herewith
;
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: McMillian, Nabeela R.
; REGISTRATION NUMBER: P-43,363
; REFERENCE/DOCKET NUMBER: ARAG:003
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
;
; INFORMATION FOR SEQ ID NO: 119:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 243 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
;
US-09-079-030-119

```

```

Query Match      91.1%; Score 1241; DB 4; Length 243;
Best Local Similarity 100.0%; Pred. No. 1.4e-98;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      25  DEBPQSPWDRVKDLATYYVDVLKDSGRDYVVSQFEGSALGKQLNLKLLDNWDSVTSTFSKL 84
      |||
Db      1  DEBPQSPWDRVKDLATYYVDVLKDSGRDYVVSQFEGSALGKQLNLKLLDNWDSVTSTFSKL 60

Qy      85  REOLGPPVTOBFNDNLSEKETEGTLQEMSKDLEEVKAKVPYLLDDPFQKKWQEMELYROKVE 144
      |||
Db      61  REOLGPPVTOBFNDNLSEKETEGTLQEMSKDLEEVKAKVPYLLDDPFQKKWQEMELYROKVE 120

Qy     145  PLRAELQEGARQKHLHQELKSLPLGEMRDRARAHVDALRTHLAPYSDELRQRLAARLEA 204
      |||
Db     121  PLRAELQEGARQKHLHQELKSLPLGEMRDRARAHVDALRTHLAPYSDELRQRLAARLEA 180

Qy     205  LKNGGARLAAYEHAKATEHLSTLSEKAKPALEDLRQGLLPVLESFKVSPFLSAALEYTKKL 264
      |||
Db     181  LKNGGARLAAYEHAKATEHLSTLSEKAKPALEDLRQGLLPVLESFKVSPFLSAALEYTKKL 240

Qy     265  NTQ 267
      |||
Db     241  NTQ 243

RESULT 8
US-08-448-606-6
: Sequence 6, Application US/08448606

```


Patent No. 5721114
GENERAL INFORMATION:
APPLICANT: Abrahams n, Lars
APPLICANT: Holmgren, Erik
APPLICANT: Kalden n, Christina
APPLICANT: Lake, Mats
APPLICANT: Mikaelsson, sa
APPLICANT: Sejlitz, Torsten
TITLE OF INVENTION: Expression System For Producing Apolipoprotein AI-M
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pollock, Vande Sande & Priddy
STREET: 1990 M Street, N.W., Suite 800
CITY: Washington
STATE: D.C.
COUNTRY: U.S.
ZIP: 20036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE: 25-AUG-1995
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/SE93/01061
FILING DATE: 09-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: SE 9203753-0
FILING DATE: 11-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Amernick, Burton A.
REGISTRATION NUMBER: 24,852
REFERENCE/DOCKET NUMBER: 0151/00121
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 331-7111
TELEFAX: (202) 293-6229
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 264 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-448-606-6

Query Match 91.1%; Score 1240.5; DB 1; Length 264;
Best Local Similarity 92.2%; Pred. No. 1.8e-98;
Matches 249; Conservative 3; Mismatches 9; Indels 9; Gaps 2;
QY 1 MKAAVLTAVL---FUTGQARHFVQDEPPQSPWRVKDLATVYVDVLKDSGRDYVSQF 57
Db 1 MKKTAIAIALAGFATVANA-----DEPPQSPWRVKDLATVYVDVLKDSGRDYVSQF 54
QY 58 EGSALGKQLNLKLLDNWDSVTSTFSKLRQLGPGVTQEFWDNLEKETEGLRQEMSKDLEEV 117
Db 55 EGSALGKQLNLKLLDNWDSVTSTFSKLRQLGPGVTQEFWDNLEKETEGLRQEMSKDLEEV 114
QY 118 KAKVQPVLDPFQKKWQEMELYRQKVEPLRAELQEGARQKLHQLQKLSPLGEMDRAR 177
Db 115 KAKVQPVLDPFQKKWQEMELYRQKVEPLRAELQEGARQKLHQLQKLSPLGEMDRAR 174
QY 178 AHVDALRTHLAPYSDELQRQAARLEALKENGARLAELYHAKATEHLSTLSEKAKPALED 237
Db 175 AHVDALRTHLAPYSDELQRQAARLEALKENGARLAELYHAKATEHLSTLSEKAKPALED 234
QY 238 LRQGLLPVLESFKVSFLSALEBYTKKLNQ 267
Db 235 LRQGLLPVLESFKVSFLSALEBYTKKLNQ 264

RESULT 9
US-08-952-796-15
Sequence 15, Application US/08952796
Patent No. 6258596
GENERAL INFORMATION:
APPLICANT: BENOIT, Patrick
APPLICANT: BRUCKERT, Eric
APPLICANT: DENEFELE, Patrice
APPLICANT: DUBERGER, Nicolas
APPLICANT: FRUCHART, Jean-Charles
APPLICANT: LUC, Gerald
APPLICANT: TURPIN, Gerard
APPLICANT: ASSMANN, Gerd
APPLICANT: FUNKE, Harald
TITLE OF INVENTION: NOVEL VARIANTS OF APOLIPOPROTEIN A-I
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, Mailstop 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/952,796
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 95/06061
FILING DATE: 22-MAY-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO FR96/00747
FILING DATE: 20-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Fehlner Esq., Paul P.
REGISTRATION NUMBER: 35,135
REFERENCE/DOCKET NUMBER: ST95031-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 200 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-952-796-15
Query Match 73.9%; Score 1006; DB 3; Length 200;
Best Local Similarity 99.5%; Pred. No. 1.5e-78;
Matches 199; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 68 LKLLDNWDSVTSTFSKLRQLGPGVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDD 127
Db 1 LKLLDNWDSVTSTFSKLRQLGPGVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDD 60
QY 128 FOKKWOEEMELYRQKVEPLRAELQEGARQKLHQLQKLSPLGEMDRARAHVDALRTHL 187
Db 61 FOKKWOEEMELYRQKVEPLRAELQEGARQKLHQLQKLSPLGEMDRARAHVDALRTHL 120
QY 188 APYSDELQRQAARLEALKENGARLAELYHAKATEHLSTLSEKAKPALEDLRQGLLPVLE 247
Db 121 APYSDELQRQAARLEALKENGARLAELYHAKATEHLSTLSEKAKPALEDLRQGLLPVLE 180
QY 248 SFKVSFLSALEBYTKKLNQ 267
Db 181 SFKVSFLSALEBYTKKLNQ 200

```

RESULT 10
US-08-292-870-1
; Sequence 1, Application US/08292870
; Patent No. 5814467
; GENERAL INFORMATION:
; APPLICANT: Curtiss, Linda K
; APPLICANT: Banks, Carole L
; APPLICANT: Bonnet, David J
; APPLICANT: Smith, Richard S
; TITLE OF INVENTION: APO AI POLYPEPTIDES, DIAGNOSTIC METHODS
; TITLE OF INVENTION: AND SYSTEMS FOR QUANTIFYING APO AI, AND THERAPEUTIC
; TITLE OF INVENTION: METHODS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of
; ADDRESSEE: Patent Counsel
; STREET: 10666 No. 5814467th Torrey Pines Road., TPC 8
; CITY: La Jolla
; STATE: California
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/292,870
; FILING DATE: 17-AUG-1994
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/534,761
; FILING DATE: 07-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/711,333
; FILING DATE: 06-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US 91/04038
; FILING DATE: 07-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: TSRI 210.1 D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 64 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-08-292-870-1

Query Match 24.1%; Score 328; DB 2; Length 64;
Best Local Similarity 100.0%; Pred. No. 3.7e-21;
Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 109 EMSKDL EEKAKVQPYLDDFKKKWQEEMLYRQKVEPLRAELQEGAROKHLEQKLSPL 166
Db 1 EMSKDL EEKAKVQPYLDDFKKKWQEEMLYRQKVEPLRAELQEGAROKHLEQKLSPL 166
QY 169 GEEM 172
Db 61 GEEM 64

RESULT 11
US-09-800-729-207
; Sequence 207, Application US/09800729
; Patent No. 6605592

```



```

; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/292,870
; FILING DATE: 17-AUG-1994
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/534,761
; FILING DATE: 07-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/711,333
; FILING DATE: 06-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US 91/04038
; FILING DATE: 07-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: TSRI 210.1 D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 44 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
;
US-08-292-870-2

```

```

Query Match      16.7%  Score 227; DB 2; Length 44;
Best Local Similarity 100.0%; Pred.No. 9.9e-13;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      107 RQEMSKDLBEEVKAKVQPYLDDFQKWKQEEEMELYRQKVEPLRAEL 150
      |||||||||||||||||||||||||||||||||||||||||||
Db      1 RQEMSKDLBEEVKAKVQPYLDDFQKWKQEEEMELYRQKVEPLRAEL 44

```

Search completed: December 21, 2004, 07:32:17
Job time : 40.7949 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2004, 06:40:58 ; Search time 32.2747 Seconds
(without alignments)
795.975 Million cell updates/sec

Title: US-09-803-918A-2

Perfect score: 1362

Sequence: 1 MKAAVLTLAVLFLTGSQARH.....SPKVSFLSAEYTKKLNTQ 267

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: PIR 79:*
- 2: piri:*
- 3: piri2:*
- 4: piri3:*
- 5: piri4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1362	100.0	267	1 LPHUA1	apolipoprotein A-I
2	1306	95.9	267	2 JS0079	apolipoprotein A-I
3	1299	95.4	267	1 A26529	apolipoprotein A-I
4	1161.5	85.3	266	1 LPDGA1	apolipoprotein A-I
5	1100.5	80.8	266	1 LPRB1Z	apolipoprotein A-I
6	1100	80.8	265	2 A46018	apolipoprotein AI
7	1093	80.2	265	2 JT0672	apolipoprotein A-I
8	1087.5	79.8	264	2 S31394	apolipoprotein A-I
9	1082	79.4	265	1 LPRB1B	apolipoprotein A-I
10	1082	79.4	265	2 A56858	apolipoprotein A-I
11	974.5	71.5	241	2 A24998	apolipoprotein A-I
12	924	67.8	231	2 JQ0704	apolipoprotein A-I
13	904.5	66.4	264	2 S22420	apolipoprotein A-I
14	898.5	66.0	262	2 JCI237	apolipoprotein A-I
15	849.5	62.4	259	2 A24700	apolipoprotein A-I
16	688.5	50.6	264	1 LPCA1	apolipoprotein A-I
17	663.5	48.7	264	2 JCS456	apolipoprotein A-I
18	644.5	47.3	246	2 A61448	apolipoprotein A-I
19	617.5	45.3	164	2 S21830	apolipoprotein A-I
20	291	21.4	439	2 S29565	apolipoprotein A-I
21	278.5	20.4	396	1 LPHUA4	apolipoprotein A-I
22	267	19.6	391	1 LPRUA4	apolipoprotein A-I
23	253	18.6	399	2 C40892	apolipoprotein A-I
24	251	18.4	391	2 B40892	apolipoprotein A-I
25	251	18.4	395	2 A40892	apolipoprotein A-I
26	249.5	18.3	401	2 A47141	apolipoprotein A-I
27	238.5	17.5	394	2 A52581	apolipoprotein A-I
28	225.5	16.6	258	2 JH0472	apolipoprotein A-I
29	209	15.3	311	2 JU0036	apolipoprotein E p

30	198	14.5	312	1 LPRTE	apolipoprotein E p
31	194.5	14.3	1547	2 T28657	blackjack protein,
32	190.5	14.0	317	2 A28792	apolipoprotein E p
33	189.5	13.9	317	2 S03185	apolipoprotein E p
34	188	13.8	317	2 S33450	apolipoprotein E -
35	184	13.5	298	2 S12635	apolipoprotein E p
36	182	13.4	329	2 JCS566	apolipoprotein E -
37	178.5	13.1	291	2 C60940	apolipoprotein E -
38	174.5	12.8	317	1 LPHUE	apolipoprotein E p
39	166.5	12.2	316	2 JCS49	apolipoprotein E p
40	160	11.7	316	2 I45996	apolipoprotein E -
41	160	11.7	316	2 S26478	apolipoprotein E -
42	147	10.8	311	2 A45951	apolipoprotein E p
43	146	10.7	513	2 S08381	keratin, 58K type
44	143.5	10.5	771	1 A33430	h-caldesmon - chic
45	137	10.1	470	2 T23512	hypothetical prote

ALIGNMENTS

RESULT 1

LPHUA1

apolipoprotein A-I precursor [validated] - human

N;Alternate names: apoA-I-2; apoA-I-4; preproapoA-I; prostacyclin stabilizing factor

C;Species: Homo sapiens (man)

C;Date: 14-Nov-1983 #sequence revision 14-Nov-1983 #text change 09-Jul-2004

C;Accession: A90947; B90947; S02373; A93465; A21147; A93519; B93519; A93472; A94010; A92

6137

R;Seilhamer, J.J.; Protter, A.A.; Frossard, P.; Levy-Wilson, B.

DNA 3, 309-317, 1984

A;Title: Isolation and DNA sequence of full-length cDNA and of the entire gene for huma

A;Reference number: A90947; MUID:85026665; PMID:6207999

A;Accession: A90947

A;Molecule type: DNA

A;Residues: 1-267 <SEI>

A;Cross-references: UNIPROT:P02647; GB:X01038; NID:g28769; PIDN:CAA25519.1; PID:g296635

A;Accession: B90947

A;Molecule type: mRNA

A;Residues: 1-267 <SE2>

A;Cross-references: GB:X01038; NID:g28769; PIDN:CAA25519.1; PID:g296635

R;Makrides, S.C.; Ruiz-Opazo, N.; Hayden, M.; Nussbaum, A.L.; Breslow, J.L.; Zannis, V.

Eur. J. Biochem. 173, 465-471, 1988

A;Title: Sequence and expression of Tangier apoA-I gene.

A;Reference number: S02373; MUID:88196137; PMID:3129297

A;Accession: S02373

A;Molecule type: DNA

A;Residues: 1-267 <NAK>

A;Cross-references: EMBL:X07496; NID:g28774; PIDN:CAA30377.1; PID:g296729

R;Shoulders, C.C.; Kornblitt, A.R.; Munro, B.S.; Baralle, F.E.

Nucleic Acids Res. 11, 2827-2837, 1983

A;Title: Gene structure of human apolipoprotein AI.

A;Reference number: A93465; MUID:83220822; PMID:6406984

A;Accession: A93465

A;Molecule type: DNA

A;Residues: 1-267 <SHO>

A;Cross-references: GB:J00098; GB:J00100; GB:J00101; GB:J03222; GB:K01518; G

R;Karthanas, S.K.; Zannis, V.I.; Breslow, J.L.

Proc. Natl. Acad. Sci. U.S.A. 80, 6147-6151, 1983

A;Title: Isolation and characterization of the human apolipoprotein A-I gene.

A;Reference number: A21147; MUID:84016011; PMID:6413973

A;Accession: A21147

A;Molecule type: DNA

A;Residues: 1-267 <KAR>

A;Cross-references: GB:J00098; GB:J03222; NID:g178765; PIDN:AAB59514.1; PID:g178768

R;Sharpe, C.R.; Sidoli, A.; Shelley, C.S.; Lucero, M.A.; Shoulders, C.C.; Baralle, F.E.

Nucleic Acids Res. 12, 3917-3932, 1984

A;Title: Human apolipoproteins AI, AII, CII and CI. cDNA sequences and mRNA abundance

A;Reference number: A93519; MUID:84221405; PMID:6328445

A;Accession: A93519

A;Molecule type: mRNA

A;Residues: 1-267 <SHA>

A;Cross-references: GB:X00566; NID:g28765; PIDN:CAA25232.1; PID:g732753

A;Accession: B93519 A;Molecule type: DNA A;Residues: 1-24 <SH2> R;Cheung, P.; Chan, L. A;Title: Nucleotide sequence of cloned cDNA of human apolipoprotein A-I. A;Reference number: A93472; MUID:83220772; PMID:6304641 A;Accession: A93472 A;Molecule type: mRNA A;Residues: 1-267 <CHE> A;Cross-references: GB:J00098; GB:J00099; GB:J00100; GB:J00101; GB:J03222; GB:K01518; GB:R;Law, S.W.; Brewer Jr., H.B. Proc. Natl. Acad. Sci. U.S.A. 81, 66-70, 1984 A;Title: Nucleotide sequence and the encoded amino acids of human apolipoprotein A-I mRNA A;Reference number: A94010; MUID:84119464; PMID:6198645 A;Accession: A94010 A;Molecule type: mRNA A;Residues: 1-267 <LAW> A;Cross-references: GB:J00099; GB:J00100; GB:J00101; GB:J03222; GB:K01518; GB:R;Zannis, V.I.; Karathanasis, S.K.; Keutmann, H.T.; Goldberger, G.; Breslow, J.L. Proc. Natl. Acad. Sci. U.S.A. 80, 2574-2578, 1983 A;Title: Intracellular and extracellular processing of human apolipoprotein A-I: secreted A;Reference number: A21118; MUID:83195100; PMID:6405383 A;Accession: A21118 A;Molecule type: mRNA A;Residues: 1-24 <ZAN> R;Brewer Jr., H.B.; Fairwell, T.; Kay, L.; Meng, M.; Ronan, R.; Law, S.; Light, J.A. Biochem. Biophys. Res. Commun. 113, 626-632, 1983 A;Title: Human plasma proapoA-I: isolation and amino-terminal sequence. A;Reference number: A90112; MUID:83256553; PMID:6409108 A;Accession: A90112 A;Molecule type: protein A;Residues: 19-27 <BRE> R;Brewer Jr., H.B.; Fairwell, T.; LaRue, A.; Ronan, R.; Houser, A.; Bronzert, T.J. Biochem. Biophys. Res. Commun. 80, 623-630, 1978 A;Title: The amino acid sequence of human APOA-I, an apolipoprotein isolated from high density lipoproteins. A;Reference number: A90209; MUID:78123731; PMID:204308 A;Accession: A90209 A;Molecule type: protein A;Residues: 25-57, 'Q', 59-169, 'QQ', 172-267 <BR2> R;Yui, Y.; Aoyama, T.; Morishita, H.; Takahashi, M.; Takatsu, Y.; Kawai, C. J. Clin. Invest. 82, 803-807, 1988 A;Title: Serum prostacyclin stabilizing factor is identical to apolipoprotein A-I (Apo A-I). A;Reference number: A30516; MUID:88331387; PMID:3047170 A;Accession: A30516 A;Molecule type: protein A;Residues: 25-56 <YUI> R;Nichols, W.C.; Dwulet, F.E.; Liepnies, J.; Benson, M.D. Biochem. Biophys. Res. Commun. 156, 762-768, 1988 A;Title: Variant apolipoprotein AI as a major constituent of a human hereditary amyloid. A;Reference number: A31582; MUID:89050104; PMID:3142462 A;Accession: A31582 A;Molecule type: protein A;Residues: 25-49, 'R', 51-85, 'D', 87-107 <NIC> A;Note: variant sequence from patient with familial amyloidotic polyneuropathy type III R;Manjunath, P.; Marcel, Y.L.; Una, J.; Seidah, N.G.; Chretien, M.; Chapdelaine, A. J. Biol. Chem. 264, 16853-16857, 1989 A;Title: Apolipoprotein A-I binds to a family of bovine seminal plasma proteins. A;Reference number: A34409; MUID:89380318; PMID:2506184 A;Accession: A34409 A;Molecule type: protein A;Residues: 25-48 <VAN> R;Stoffel, W.; Binczek, E. Biol. Chem. Hoppe-Seyler 369, 1055-1063, 1988 A;Title: Structural requirements of human preproapolipoprotein AI for translocation and A;Reference number: S02737; MUID:89149957; PMID:3328490 A;Accession: S02737 A;Status: not compared with conceptual translation A;Molecule type: mRNA A;Residues: 1-6, 'AV', 9, 'LV', 12-29 <STO> A;Note: part of this sequence, including the amino end of the mature protein, was confirmed by sequencing. R;Stoffel, W.; Binczek, E. Biol. Chem. Hoppe-Seyler 372, 481-488, 1991	A;Title: Transient expression of wild type and mutant human apolipoprotein AI in COS cells. A;Reference number: S16197; MUID:92029676; PMID:1930731 A;Contents: annotation; extension of studies in reference S02737 R;Stoffel, W.; Kruger, E.; Deutzmann, R. Hoppe-Seyler's Z. Physiol. Chem. 364, 227-237, 1983 A;Title: Cell-free translation of human liver apolipoprotein AI and AII mRNA processing A;Reference number: A19913; MUID:83236195; PMID:6407957 A;Accession: B19913 A;Molecule type: protein A;Residues: 1-6, 'X', 8-13, 'XXX', 17-18, 'XX', 21, 'X', 23-25, 'X', 27-29 <ST2> R;Ehnholm, C.; Bozas, S.E.; Tenkanen, H.; Kirsbaum, L.; Metso, J.; Murphy, B.; Walker, B. Biochim. Biophys. Acta 1086, 255-260, 1991 A;Title: The apolipoprotein A-I binding protein of placenta and the SP-40,40 protein of A;Reference number: A56815; MUID:92075698; PMID:1742316 A;Accession: A56815 A;Molecule type: protein A;Residues: 25-31, 'P', 33 <ERN> A;Experimental source: serum A;Note: sequence extracted from NCBI backbone (NCBIP:69759) A;Note: 32-Trp was also found R;Kunitake, S.T.; Carilli, C.T.; Lau, K.; Protter, A.A.; Naya-Vigne, J.; Kane, J.P. Biochemistry 33, 1988-1993, 1994 A;Title: Identification of proteins associated with apolipoprotein A-I-containing lipoproteins. A;Reference number: A54223; MUID:94162201; PMID:8117655 A;Accession: A54223 A;Molecule type: protein A;Residues: 25-39 <KUN> R;Mogulievsky, N.; Roobol, C.; Loriau, R.; Guillaume, J.P.; Jacobs, P.; Cravador, A.; He DNA 8, 429-436, 1989 A;Title: Production of human recombinant proapolipoprotein A-I in Escherichia coli: purification and characterization. A;Reference number: I39476; MUID:89377481; PMID:2673706 A;Accession: I39476 A;Molecule type: mRNA A;Residues: 19-267 <RES> A;Cross-references: GB:M29068; NID:gl78774; PIDN:AAA51747.1; PID:gl78775 R;Higuchi, K.; Law, S.W.; Hoeg, J.M.; Schumacher, U.K.; Meglin, N.; Brewer, H.B. J. Biol. Chem. 263, 18530-18536, 1988 A;Title: Tissue-specific expression of apolipoprotein A-I (ApoA-I) is regulated by the 5' noncoding region. A;Reference number: I39475; MUID:89050400; PMID:3142880 A;Accession: I39475 A;Molecule type: DNA A;Residues: 1-14 <R2> A;Cross-references: GB:J04066; NID:gl78763; PIDN:AAA51746.1; PID:g553183 R;Breslow, J.L. Annu. Rev. Biochem. 54, 699-727, 1985 A;Title: Human apolipoprotein molecular biology and genetic variation. A;Reference number: A30042; MUID:85278004; PMID:3896129 A;Contents: annotation; review of sequences, variants and gene location R;Hoeg, J.M.; Meng, M.S.; Ronan, R.; Fairwell, T.; Brewer Jr., H.B. J. Biol. Chem. 261, 3911-3914, 1986 A;Title: Human apolipoprotein A-I. Post-translational modification by fatty acid acylation. A;Reference number: A32577; MUID:86140194; PMID:3005308 A;Contents: annotation; acylation with palmitate A;Note: an undetermined serine or threonine is acylated by fatty acid, the acylating fatty acid is not specified. R;Law, S.W.; Brewer, H.B. J. Biol. Chem. 260, 12810-12814, 1985 A;Title: Tangier disease: The complete mRNA sequence encoding for preproapoA-I. A;Reference number: I55236; MUID:86008382; PMID:2995392 A;Accession: I55236 A;Status: preliminary; translated from GB/EMBL/DBJ A;Molecule type: mRNA A;Residues: 1-143, 'D', 145-267 <RE3> A;Cross-references: GB:M11791; NID:gl78776; PIDN:AAA35545.1; PID:gl78777 C;Comment: Apolipoprotein A-I precursor is synthesized in the liver and small intestine C;Genetics: HDL in plasma. A;Gene: GDB:APOA1 A;Cross-references: GDB:119684; OMIM:107680 A;Map position: 11q23.3-11q23.3 A;Introns: 15/1; 67/2 C;Function: A;Description: participates in the reverse transport of cholesterol from tissues to the sterol acyltransferase (LCAT); noncovalently binds and stabilizes prostacyclin (PGI-2)
---	--

C;Superfamily: apolipoprotein A-I	
C;Keywords: atherosclerosis; cholesterol metabolism; HDL; intestine; lipid binding; lipid	
F;1-18/Domain: signal sequence #status experimental <SIG>	
F;19-24/Domain: propeptide #status experimental <PRO>	
F;25-267/Product: apolipoprotein A-I #status experimental <MAT>	
Query Match 100.0%; Score 1362; DB 1; Length 267;	
Best Local Similarity 100.0%; Pred. No. 3.3e-69;	
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 MKAAVLTAVLFLTGSQARHFHQDEPPQSPMDRVKDLATVYVDVLKDSGRDVSQFEGS 60
Db	1 MKAAVLTAVLFLTGSQARHFHQDEPPQSPMDRVKDLATVYVDVLKDSGRDVSQFEGS 60
QY	61 ALGKQLNLKLLDNWDSVTSTFSKLRQLGPGVPTQEFWDNLEKETEGRLQEMSKDLEEVKAK 120
Db	61 ALGKQLNLKLLDNWDSVTSTFSKLRQLGPGVPTQEFWDNLEKETEGRLQEMSKDLEEVKAK 120
QY	121 VQPYLDDFOKKQWQEMELRYQKVEPLRAELQEGARQKLHQLQELKSPGLGEEMDRARAHV 180
Db	121 VQPYLDDFOKKQWQEMELRYQKVEPLRAELQEGARQKLHQLQELKSPGLGEEMDRARAHV 180
QY	181 DALRTHLAPYSDELQRLAARLEALKENGCGARLAEYHAKATEHLSTLSEKAKPALEDLRQ 240
Db	181 DALRTHLAPYSDELQRLAARLEALKENGCGARLAEYHAKATEHLSTLSEKAKPALEDLRQ 240
QY	241 GLLPVLESFKVSFLSALEYTKKLNTQ 267
Db	241 GLLPVLESFKVSFLSALEYTKKLNTQ 267
RESULT 2	
JS0079	
apolipoprotein A-I precursor - baboon	
C;Species: Papio sp. (baboon)	
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 13-Jun-1997	
C;Accession: JS0079	
R;Hixon, J.E.; Borenstein, S.; Cox, L.A.; Rainwater, D.L.; VandeBerg, J.L.	
Gene 74, 483-490, 1988	
A;Title: The baboon gene for apolipoprotein A-I: Characterization of a cDNA clone and id	
A;Reference number: JS0079; MUID:89232739; PMID:2907746	
A;Accession: JS0079	
A;Molecule type: mRNA	
A;Residues: 1-267 <HIX>	
A;Experimental source: liver	
A;Comment: This protein is the principal protein component of high density lipoprotein p	
C;Comment: This protein serves as a cofactor for lecithin cholesterol acyltransferase co	
C;Comment: This protein contains a region of repeated amino acids which form amphipathic	
C;Genetics:	
A;Gene: apoA1	
C;Superfamily: apolipoprotein A-I	
C;Keywords: HDL; lipid binding; lipoprotein	
F;1-18/Domain: signal sequence #status predicted <SIG>	
F;19-267/Product: apolipoprotein A-I #status predicted <IAI>	
F;123-144,145-166,167-188,189-210,211-232,233-254/Region: tandem repeats	
Query Match 95.9%; Score 1306; DB 2; Length 267;	
Best Local Similarity 95.1%; Pred. No. 4.4e-66;	
Matches 254; Conservative 6; Mismatches 7; Indels 0; Gaps 0;	
QY	1 MKAAVLTAVLFLTGSQARHFHQDEPPQSPMDRVKDLATVYVDVLKDSGRDVSQFEGS 60
Db	1 MKATVLTAVLFLTGSQARHFHQDEPPQTPMDRVKDLVTVYVEALKDSGKDVVSQFEGS 60
QY	61 ALGKQLNLKLLDNWDSVTSTFSKLRQLGPGVPTQEFWDNLEKETEGRLQEMSKDLEEVKAK 120
Db	61 ALGKQLNLKLLDNWDSVTSTFSKLRQLGPGVPTQEFWDNLEKETEGRLQEMSKDLEEVKAK 120
QY	121 VQPYLDDFOKKQWQEMELRYQKVEPLRAELQEGARQKLHQLQELKSPGLGEEMDRARAHV 180
Db	121 VQPYLDDFOKKQWQEMELRYQKVEPLRAELHEGTROKLHQLHKLKSPGLGEEMDRARAHV 180
QY	181 DALRTHLAPYSDELQRLAARLEALKENGCGARLAEYHAKATEHLSTLSEKAKPALEDLRQ 240

Db	181 DALRTHLAPYSDELQRLAARLEALKENGCGARLAEYHAKASEHLSTLSEKAKPALEDLRQ 240
QY	241 GLLPVLESFKVSFLSALEYTKKLNTQ 267
Db	241 GLLPVLESFKVSFLSALEYTKKLSTQ 267
RESULT 3	
A26529	
apolipoprotein A-I precursor - crab-eating macaque	
C;Species: Macaca fascicularis (crab-eating macaque)	
C;Date: 30-Sep-1989 #sequence_revision 19-Oct-1995 #text_change 09-Jul-2004	
C;Accession: A26529; A26627; S23135; A57766	
R;Polltes, H.G.; Melchior, G.W.; Castle, C.K.; Marotti, K.R.	
Gene 49, 103-110, 1986	
A;Title: The primary structure of cynomolgus monkey apolipoprotein A-1 deduced from the	
A;Reference number: A26529; MUID:87191989; PMID:3106152	
A;Accession: A26529	
A;Molecule type: mRNA	
A;Residues: 1-267 <POL>	
A;Cross-references: UNIPROT:P15568; GB:M15411; NID:G342074; PIDN:AAA36834.1; PID:G34207	
R;Herbert, P.N.; Bauserman, L.L.; Lynch, K.M.; Saritelli, A.L.; Kantor, M.A.; Nicolosi	
Biochemistry 26, 1457-1463, 1987	
A;Title: Homologues of the human C and A apolipoproteins in the Macaca fascicularis (cy	
A;Reference number: A26627; MUID:87185451; PMID:3105581	
A;Accession: A26627	
A;Molecule type: protein	
A;Residues: 25-48 <HER>	
R;Murray, R.W.; Marotti, K.R.	
Biochim. Biophys. Acta 1131, 207-210, 1992	
A;Title: Nucleotide sequence of the cynomolgus monkey apolipoprotein A-I gene and corre	
A;Reference number: S23135; MUID:92305062; PMID:1610902	
A;Accession: S23135	
A;Molecule type: DNA	
A;Residues: 1-12, 'L', 14-267 <MUR>	
A;Cross-references: GB:M83242; NID:G342070; PIDN:AAA36832.1; PID:G342071	
R;Sorci-Thomas, M.; Kearns, M.W.	
J. Biol. Chem. 266, 18045-18050, 1991	
A;Title: Transcriptional regulation of the apolipoprotein A-I gene.	
A;Reference number: A57766; MUID:92011532; PMID:1917942	
A;Accession: A57766	
A;Molecule type: DNA	
A;Residues: 1-10 <RES>	
A;Cross-references: GB:M69223; NID:G342066; PIDN:AAA36831.1; PID:G553820	
C;Comment: The precursor is synthesized in the liver and small intestine. The propeptid	
C;Comment: ApoA-I makes up about 70% of the protein of the high density lipoproteins (H	
Y promoting cholesterol efflux from tissues and by acting as a cofactor for the lecithin	
C;Genetics:	
A;Introns: 15/1; 67/2	
C;Superfamily: apolipoprotein A-I	
C;Keywords: cholesterol metabolism; HDL; intestine; lipid transport; plasma; tandem rep	
F;1-18/Domain: signal sequence #status predicted <SIG>	
F;19-24/Domain: propeptide #status predicted <PPR>	
F;25-267/Product: apolipoprotein A-I #status predicted <MAT>	
Query Match 95.4%; Score 1299; DB 1; Length 267;	
Best Local Similarity 94.8%; Pred. No. 1.1e-65;	
Matches 253; Conservative 6; Mismatches 8; Indels 0; Gaps 0;	
QY	1 MKAAVLTAVLFLTGSQARHFHQDEPPQSPMDRVKDLATVYVDVLKDSGRDVSQFEGS 60
Db	1 MKATVLTAVLFLTGSQARHFHQDEPPQTPMDRVKDLVTVYVEALKDSGKDVVSQFEGS 60
QY	61 ALGKQLNLKLLDNWDSVTSTFSKLRQLGPGVPTQEFWDNLEKETEGRLQEMSKDLEEVKAK 120
Db	61 ALGKQLNLKLLDNWDSVTSTFSKLRQLGPGVPTQEFWDNLEKETEGRLQEMSKDLEEVKAK 120
QY	121 VQPYLDDFOKKQWQEMELRYQKVEPLRAELQEGARQKLHQLQELKSPGLGEEMDRARAHV 180
Db	121 VQPYLDDFOKKQWQEMELRYQKVEPLRAELHEGTROKLHQLHKLKSPGLGEEMDRARAHV 180
QY	181 DALRTHLAPYSDELQRLAARLEALKENGCGARLAEYHAKATEHLSTLSEKAKPALEDLRQ 240

Db 181 DALRTHLAPYSDLRQLAARLEALKENGARLAETHAKASEHLSLTSEKAPALEDLRQ 240
QY 241 GLLPVLESFKVSLSALEYTKKLNQ 267
Db 241 GLLPVLESFKVSLSALEYTKKLSQ 267

RESULT 4

LPDGA1
apolipoprotein A-I precursor - dog
C:Species: Canis lupus familiaris (dog)
C:Date: 17-Dec-1982 #sequence_revision 14-Jul-1994 #text_change 09-Jul-2004
C:Accession: A60940; A03092; R61418
R;Luo, C.C.; Li, W.H.; Chan, L.
J. Lipid Res. 30, 1735-1746, 1989
A:Title: Structure and expression of dog apolipoprotein A-I, E, and C-I mRNAs: implications for the evolution of the apolipoprotein A-I gene family
A:Reference number: A60940; MUID:90132271; PMID:2515239
A:Accession: A60940
A:Molecule type: mRNA
A:Residues: 1-266 <LUO>
A:Cross-references: UNIPROT:P02648
R;Chung, H.; Randolph, A.; Reardon, I.; Heinrikson, R.L.
J. Biol. Chem. 257, 2961-2967, 1982
A:Title: The covalent structure of apolipoprotein A-I from canine high density lipoproteins
A:Reference number: A03092; MUID:82142425; PMID:6801039
A:Accession: A03092
A:Molecule type: protein
A:Residues: 25-167, 'G', 169-201, 'Q', 203-234, 'Q', 236-266 <CHU>
R;Nakai, T.; Whayne, T.F.; Tang, J.
FEBS Lett. 64, 409-411, 1976
A:Title: The amino- and carboxyl-terminal sequences of canine apolipoprotein A-I.
A:Reference number: A61418; MUID:76210910; PMID:179987
A:Accession: A61418
A:Molecule type: protein
A:Residues: 25-56, 'Z', 261-262, 'A' <NAK>
C:Superfamily: apolipoprotein A-I
C:Keywords: atherosclerosis; cholesterol metabolism; HDL; intestine; lipid transport; lipoproteins
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-24/Domain: propeptide #status predicted <PRO>
F:25-266/Product: apolipoprotein A-I #status experimental <MAT>

Query Match 85.3%; Score 1161.5; DB 1; Length 266;
Best Local Similarity 85.0%; Pred. No. 4.8e-58;
Matches 227; Conservative 19; Mismatches 20; Indels 1; Gaps 1;
QY 1 MKAALVTLAVLFTGSGARHFHQDEPPQSPMDRVKDLATVYVDVLKSGRDYVSQFEGS 60
Db 1 MKAALVTLAVLFTGSGARHFHQDE-PPQSPMDRVKDLATVYVDVAVKDSGRDYVAQFEAS 59
QY 61 ALGKQLNLKLLDNWDSVSTFSKLRQLGPGVTQEFWDNLEKETEGRLQEMSKDLEEVKAK 120
Db 60 ALGKQLNLKLLDNWDSLSSTVTVKLRQIGPGVTQEFWDNLEKETEVLRQEMSKDLEEVKQK 119
QY 121 VQPYLDDFQKKWQEEEMELYRQKVEPLRAELQEGAROKLHELQKLSPLGEMMRDRAAHV 180
Db 120 VQPYLDDFQKKWQEEVELRQKVAPLGSELREGAROKLQELQKLSPLAEELRDRAHVV 179
QY 181 DALRTHLAPYSDLRQLAARLEALKENGARLAETHAKATEHLSLTSEKAPALEDLRQ 240
Db 180 DALRAQLAPYSDLRERLAARLEALKEGGASLAETHAKASEHLSLTSEKAPALEDLRQ 239
QY 241 GLLPVLESFKVSLSALEYTKKLNQ 267
Db 240 GLLPVLESFKVSLAAIDEATKLNQ 266

RESULT 5

LPBBLZ
apolipoprotein A-I precursor (clone 2Zap AI) - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 09-Jul-2004
C:Accession: S06064

R;Paraskevopoulos, T.B.; Kritis, A.; Zannis, V.
submitted to the EMBL Data Library, July 1989
A:Reference number: S06064
A:Accession: S06064
A:Molecule type: mRNA
A:Residues: 1-266 <PAR>
A:Cross-references: UNIPROT:P09809; EMBL:X15908; NID:G1457; PIDN:CAA34024.1; PID:G1458
C:Comment: This protein is synthesized in the small intestine.
C:Comment: This protein is a major component of the high density lipoproteins in plasma.
C:Superfamily: apolipoprotein A-I
C:Keywords: cholesterol metabolism; HDL; intestine; lipid transport; plasma; tandem repeat
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-24/Domain: propeptide #status predicted <PRO>
F:25-266/Product: apolipoprotein A-I #status experimental <MAT>

Query Match 80.8%; Score 1100.5; DB 1; Length 266;
Best Local Similarity 80.1%; Pred. No. 1.2e-54;
Matches 214; Conservative 25; Mismatches 27; Indels 1; Gaps 1;
QY 1 MKAALVTLAVLFTGSGARHFHQDEPPQSPMDRVKDLATVYVDVLKSGRDYVSQFEGS 60
Db 1 MKAALVTLAVLFTGSGARHFHQDE-PPSSWDKIKDFATVYVDIVKSGRGYVAQFEAS 59
QY 61 ALGKQLNLKLLDNWDSVSTFSKLRQLGPGVTQEFWDNLEKETEGRLQEMSKDLEEVKAK 120
Db 60 AFSGKQLNLKLLDNWDSLSSTVSKLQEQLGPGVTQEFWDNLEKETEGRLREMNKDLQEVROK 119
QY 121 VQPYLDDFQKKWQEEEMELYRQKVEPLRAELQEGAROKLHELQKLSPLGEMMRDRAAHV 180
Db 120 VQPYLDEFQKKWQEEVERQKVEPLGAELRESARQKLTQELQKLSPLAEELRDSARTHV 179
QY 181 DALRTHLAPYSDLRQLAARLEALKENGARLAETHAKATEHLSLTSEKAPALEDLRQ 240
Db 180 DTLRTKLAPYSELQORLAARLESKEGGASLABYQAKAREHLSVLSEKAPALEDLRQ 239
QY 241 GLLPVLESFKVSLSALEYTKKLNQ 267
Db 240 GLLPVLESFKASQVNLDEATKLNQ 266
RESULT 6
A46018
apolipoprotein AI - pig
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C:Accession: A46018
R;Birchbauer, A.; Knipping, G.; Juritsch, B.; Aschauer, H.; Zechner, R.
Genomics 15, 643-652, 1993
A:Title: Characterization of the apolipoprotein AI and CIII genes in the domestic pig.
A:Reference number: A46018; MUID:93224154; PMID:8468059
A:Accession: A46018
A:Status: preliminary
A:Molecule type: DNA; protein
A:Residues: 1-265 <BIR>
A:Cross-references: UNIPROT:P18648; GB:I00626; NID:G164358; PIDN:AAA30992.1; PID:G164358;
A:Note: sequence extracted from NCBI backbone (NCBIN:129509, NCBIPI:129511)
C:Superfamily: apolipoprotein A-I

Query Match 80.8%; Score 1100; DB 2; Length 265;
Best Local Similarity 80.9%; Pred. No. 1.3e-54;
Matches 216; Conservative 19; Mismatches 30; Indels 2; Gaps 2;
QY 1 MKAALVTLAVLFTGSGARHFHQDEPPQSPMDRVKDLATVYVDVLKSGRDYVSQFEGS 60
Db 1 MKAALVTLAVLFTGSGARHFHQDD-PPQSPMDRVKDLATVYVDIVKSGRGYVAQFEAS 59
QY 61 ALGKQLNLKLLDNWDSVSTFSKLRQLGPGVTQEFWDNLEKETEGRLQEMSKDLEEVKAK 120
Db 60 ALGKHLNLKLLDNWDSLSSTFTVKRQELGPGVTQEFWDNLEKETEARLQEMSKDLEEVKCK 119
QY 121 VQPYLDDFQKKWQEEEMELYRQKVEPLRAELQEGAROKLHELQKLSPLGEMMRDRAAHV 180
Db 120 VQPYLDDFQKKWQEEEMELYRQKVAFLGAELFREGARQKQVQLQKLSPLAEELRDLRAHV 179


```
Best Local Similarity 79.4%; Pred. No. 1.3e-53;
Matches 212; Conservative 26; Mismatches 27; Indels 2; Gaps 2;

QY 1 MKAAVLTAVLFLTGSAQRHFQWQDEPPQSPMDRVKDLATVYVDVKDSGRDYVSQFEGS 60
DB 1 MKAAVLTAVLFLTGSAQRHFQWQDE-PRSSWDKIKDFATVYVDVTKDSGREYVAQFEAS 59

QY 61 ALGKQLNLKLLDNWDSVTSFSLREQLGPVTQEFWDNLEKETEGRLQRQMSKDLREVKAK 120
DB 60 AFGKQLNLKLLDNWDSVTSVKLQELGPVTQEFWDNLEKETEGRLQRQMSKDLREVKAK 119

QY 121 VOPYLDDFOKKWOEEMELYRQKVEPLRAELQEGARQKLHELOEKLSPGLGEEMRDRARAHV 180
DB 120 VQPFDFEFQKWOEEVRYRQKVEPLRAELQEGARQKLHELOEKLSPGLGEEMRDRARAHV 179

QY 181 DALRTHLAPYSDELRLQRLAARLEALKENGARLAELAEYHAKATEHLSTLSEKAKPALEDLRQ 240
DB 180 DTLRTKLAPYSNEL-QRLAARLESIKEGGAKLAELAEYHAKATEHLSTLSEKAKPALEDLRQ 238

QY 241 GLLPVLESFKVSPFLSALEYTKKLNTQ 267
DB 239 GLLPVLESFKASQNVVDATKKLNTQ 265

RESULT 10
A56858
apolipoprotein A-I precursor - bovine
C:Species: Bos primigenius taurus (cattle)
C>Date: 05-Jan-1996 #sequence_revision 23-Aug-1997 #text_change 09-Jul-2004
C:Accession: I45853; A56858; A34649
R:O'Huigin, C.; Chan, L.; Li, W.
Mol. Biol. Evol. 7, 327-339, 1990
A>Title: Cloning and sequencing of bovine apolipoprotein A-I cDNA and molecular evolution
A:Reference number: I45853; MUID:90348478; PMID:2117227
A:Accession: I45853
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-265 <OXH>
A:Cross-references: UNIPROT:P15497; GB:M35870; NID:g162677; PIDN:AAA30381.1; PID:g162678
R:Sparrow, D.A.; Lee, B.R.; Laplaud, P.M.; Laplaud, P.M.; Laplaud, M.J.; Gd
Biochim. Biophys. Acta 1123, 145-150, 1992
A>Title: Plasma lipid transport in the prerinant calf, Bos spp: primary structure of b
A:Reference number: A56858; MUID:92153895; PMID:1739745
A:Accession: A56858
A>Status: preliminary
A:Molecule type: protein
A:Residues: 19-184, 'QL', 187-265 <SPA>
A:Experimental source: Friesian-Holstein male calves aged 2-4 weeks
A>Note: sequence extracted from NCBI backbone (NCBIP:83520)
R:Auboin, S.; Sparrow, D.A.; Beaubatie, L.; Bauchart, D.; Sparrow, J.T.; Laplaud, P.M.
Biochem. Biophys. Res. Commun. 166, 833-839, 1990
A>Title: Characterization and amino-terminal sequence of apolipoprotein AI from plasma h
A:Reference number: A34649; MUID:90147795; PMID:2105728
A:Accession: A34649
A:Molecule type: protein
A:Residues: 25-70 <AUB>
A:Experimental source: Friesian-Holstein male calves aged 2-4 weeks
C:Superfamily: apolipoprotein A-I
C:Keywords: lipid binding; lipoprotein
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-265/Product: apolipoprotein A-I #status experimental <MAT>

Query Match 79.4%; Score 1082; DB 2; Length 265;
Best Local Similarity 78.7%; Pred. No. 1.3e-53;
Matches 210; Conservative 24; Mismatches 31; Indels 2; Gaps 2;

QY 1 MKAAVLTAVLFLTGSAQRHFQWQDEPPQSPMDRVKDLATVYVDVKDSGRDYVSQFEGS 60
DB 1 MKAAVLTAVLFLTGSAQRHFQWQDD-PQSSWDRVDKDFATVYVEAIKDSGRDYVAQFEAS 59

QY 61 ALGKQLNLKLLDNWDSVTSFSLREQLGPVTQEFWDNLEKETEGRLQRQMSKDLREVKAK 120
DB 60 ALGKQLNLKLLDNWDTLASTLSKVRQLGPVTQEFWDNLEKETASLRQEMHKDLREVKQK 119
```

```
QY 121 VOPYLDDFOKKWOEEMELYRQKVEPLRAELQEGARQKLHELOEKLSPGLGEEMRDRARAHV 180
DB 120 VOPYLDDFOKKWOEEMELYRQKVEPLRAELQEGARQKLHELOEKLSPGLGEEMRDRARAHV 179

QY 181 DALRTHLAPYSDELRLQRLAARLEALKENGARLAELAEYHAKATEHLSTLSEKAKPALEDLRQ 240
DB 180 ETLRQHVAPYSDDLQRLAARLEALKENGARLAELAEYHAKATEHLSTLSEKAKPALEDLRQ 238

QY 241 GLLPVLESFKVSPFLSALEYTKKLNTQ 267
DB 239 GLLPVLESIKVSIILAAIDEASKKLNAQ 265

RESULT 11
A24998
apolipoprotein A-I - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C>Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 31-Dec-1993
C:Accession: A24998
R:Yang, C.; Yang, T.; Pownall, H.J.; Gatto Jr., A.M.
Eur. J. Biochem. 160, 427-431, 1986
A>Title: The primary structure of apolipoprotein A-I from rabbit high-density lipoprotein
A:Reference number: A24998; MUID:87030294; PMID:3095115
A:Accession: A24998
A:Molecule type: protein
A:Residues: 1-241 <YAN>
C:Superfamily: apolipoprotein A-I
C:Keywords: HDL; lipid binding; lipoprotein

Query Match 71.5%; Score 974.5; DB 2; Length 241;
Best Local Similarity 78.5%; Pred. No. 1.1e-47;
Matches 190; Conservative 23; Mismatches 28; Indels 1; Gaps 1;

QY 26 EPPQSPMDRVKDLATVYVDVKDSGRDYVSQFEGSALGKQLNLKLLDNWDSVTSFSLR 85
DB 1 DEPRSSWDKIKDFATVYVDV-KDSGREYVAQFEASAFGKQLNLKLLDNWDSVTSVKLQ 59

QY 86 EQLGPVTQEFWDNLEKETEGRLQRQMSKDLREVKAKVQVYLDLDFQKKWOEEMELYRQKVEP 145
DB 60 EQLGPVTQEFWDNLEKETEGRLQRQMSKDLREVKAKVQVYLDLDFQKKWOEEMELYRQKVEP 119

QY 146 LRAELQEGARQKLHELOEKLSPGLGEEMRDRARAHVDALRTHLAPYSDELRLQRLAARLEAL 205
DB 120 LGVELRESARQKLTELQEKLSPLAELRDSARTHTVDTLTKLAPYSQELQRLAARLESI 179

QY 206 KENGARLAELAEYHAKATEHLSTLSEKAKPALEDLRGLLPVLESFKVSPFLSALEYTKKL 265
DB 180 KEGGASLAELAEYQAKAREHLSVLSEKARPALEDLRGLLPVLESFKASQNVLDATKKL 239

QY 266 TQ 267
DB 240 TQ 241
```

```
RESULT 12
JQ0704
apolipoprotein A-I - pig (fragment)
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 28-Oct-1994
C:Accession: JQ0704
R:Weiler-Guettler, H.; Sommerfeldt, M.; Papandriopoulou, A.; Mischek, U.; Bonitz, D.; I
J. Neurochem. 54, 444-450, 1990
A>Title: Synthesis of apolipoprotein A-1 in pig brain microvascular endothelial cells.
A:Reference number: JQ0704; MUID:90132667; PMID:2105375
A:Accession: JQ0704
A:Molecule type: mRNA
A:Residues: 1-231 <WEI>
A>Note: the authors translated the codon CAG for residue 124 as His and GAC for residue
C:Superfamily: apolipoprotein A-I
C:Keywords: cholesterol metabolism; HDL; lipid binding; lipid transport; lipoprotein; p

Query Match 67.8%; Score 924; DB 2; Length 231;
```

Best Local Similarity 79.0%; Pred. No. 6.9e-45;
Matches 184; Conservative 18; Mismatches 29; Indels 2; Gaps 2;

QY 35 VKDLATVYVDVLDKSGRDYVSQFEGSALGKQLNLKLLDNWDSVTSTFSKLEQLRGQVPTQE 94
Db 1 VKDFATVYVDVLDKSGRDYVQAQFEASALGKHLNLKLLDNWDSLGSTFTKVRQELGSPVTOE 60

QY 95 FWNLEKETEGLRQEMSKOLEEVKAKVQPYLDLFQKKWQOEMELYRQKVEPLRAELQEGA 154
Db 61 FWNLEKETELAQKMSKOLEEVKKVQPYLDLFQNKWQOEMETYYRQKM-PLGAEEFREGA 119

QY 155 RQKLEHLEQKLSPLGEMDRDRAHVDAURTHAPYSDELORLARLEALKENGARLA 214
Db 120 RQKVQLEQKLSPLAELDRLARHVAALQHVAPYSDDLQRMAARFEALKE-GGDSLA 178

QY 215 EYHAKATEHLSTLSEKAPALEDLROGLLPVLESFKVSFLSALEYTKKLNTQ 267
Db 179 EYQAKAQEQKALGEKAKPALEDLROGLLPVLENLKVSIILAAIDEASKKLNAQ 231

RESULT 13
S22420
apolipoprotein A-I precursor - mouse
C:Species: Mus musculus (house mouse)
C>Date: 13-Jan-1995 #sequence revision 13-Jan-1995 #text_change 09-Jul-2004
C:Accession: S22420; S22421; A44364
R:Stoffel, W.; Mueller, R.; Binczek, E.; Hofmann, K.
Biol. Chem. Hoppe-Seyler 373, 187-193, 1992
A:Title: Mouse apolipoprotein A-I. cDNA-derived primary structure, gene organisation and
A:Reference number: S22420; MUID:92281682; PMID:1596360
A:Accession: S22420
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-264 <STO>
A:Cross-references: UNIPROT:Q00623; EMBL:X64262; NID:g50014; PIDN:CAA45560.1; PID:g50015
A:Accession: S22421
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-264 <ST2>
A:Cross-references: EMBL:X64263; NID:g50020; PIDN:CAA45561.1; PID:g50021
R:Januzzi, J.F.; Azrolan, N.; O'Connell, A.; Aalto-Setälä, K.; Breslow, J.L.
Genomics 14, 1081-1088, 1992
A:Title: Characterization of the mouse apolipoprotein Apoa-1/Apoc-3 gene locus: genomic,
A:Reference number: A44364; MUID:93122774; PMID:1478650
A:Accession: A44364
A:Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-264 <JAN>
A:Note: sequence extracted from NCBI backbone (NCBIN:122400, NCBIPI:122407)
C:Genetics:
A:Introns: 15/1; 66/2
C:Superfamily: apolipoprotein A-I

Query Match 66.4%; Score 904.5; DB 2; Length 264;
Best Local Similarity 65.2%; Pred. No. 9.7e-44;
Matches 174; Conservative 41; Mismatches 49; Indels 3; Gaps 2;

QY 1 MKAAVLTAVLFLTGSQARHFQWDEPPQSPMDRVKDLATVYVDVLDKSGRDYVSQFEGS 60
Db 1 MKAAVLAVALVFLTGSQAHVWQODE-POSQMDKVDFANVYVDVAVKDSGRDYVSQFESS 59

QY 61 ALGKQLNLKLLDNWDSVTSTFSKLRQLGQVPTQEFWDNLEKETEGRLQKQMSKOLEEVKAK 120
Db 60 SLGQQLNLNLLENWDTLGSVLSQQLERLGLTRDFWNLLEKETDWDVRQEMNKOLEEVKQK 119

QY 121 VQPYLDDFOKKWQOEMELYRQKVEPLRAELQEGARQKLHELOKLSPLGEMDRDRAHV 180
Db 120 VQPYLDFQKKWQOEMELYRQKVEPLRAELQESARQKLQELQRLSPVAEEFDRMRTHV 179

QY 181 DALRTHLAPYSDELORLARLEALKENGARLAELYHAKATEHLSTLSEKAPALEDLRQ 240
Db 180 DSLRTQLAPHSEQMRESLAORLAELKSN--PTLNEVHTRAKTHLTKLGEKARPALEDLRH 237

QY 241 GLLPVLESFKVSFLSALE 258
Db 238 SLMPMLETLKTKAQSVIE 255

RESULT 15
A24700
apolipoprotein A-I precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 30-Jun-1988 #sequence revision 30-Jun-1988 #text_change 09-Jul-2004
C:Accession: A24700; S00298; A05314
R:Haddad, I.A.; Ordovas, J.M.; Fitzpatrick, T.; Karathanasis, S.K.
J. Biol. Chem. 261, 13268-13277, 1986
A:Title: Linkage, evolution, and expression of the rat apolipoprotein A-I, C-III, and A
A:Reference number: A92558; MUID:87008540; PMID:3020028
A:Accession: A24700
A:Molecule type: DNA
A:Residues: 1-259 <HAD>
A:Cross-references: UNIPROT:P04639; EMBL:J02597; NID:g202935; PIDN:AAA40745.1; PID:g202
R:Poncin, J.B.; Martial, J.A.; Gielen, J.E.
Eur. J. Biochem. 140, 493-498, 1984
A:Title: Cloning and structure analysis of the rat apolipoprotein A-I cDNA.
A:Reference number: S00298; MUID:84207987; PMID:6426556
A:Accession: S00298
A:Molecule type: mRNA
A:Residues: 1-259 <PON>

QY 241 GLLPVLESFKVSFLSALEYTKKLNTQ 267
Db 238 SLMPMLETLKTKAQSVIDKASETLTAQ 264

RESULT 14
JC1237
apolipoprotein A-I precursor - mouse
C:Species: Mus musculus (house mouse)
C>Date: 30-Sep-1993 #sequence revision 30-Sep-1993 #text_change 31-Dec-1993
C:Accession: JC1237
R:Boyle, T.P.; Marotti, K.R.
Gene 117, 243-247, 1992
A:Title: Structure of the murine gene encoding apolipoprotein A-I.
A:Reference number: JC1237; MUID:92347700; PMID:1639271
A:Accession: JC1237
A:Molecule type: DNA
A:Residues: 1-262 <BOY>
A:Cross-references: GB:M77801
C:Genetics:
A:Gene: ApoA-I
A:Introns: 15/1; 66/2
C:Superfamily: apolipoprotein A-I
C:Keywords: cholesterol metabolism; HDL; intestine; lipid binding; lipid transport; lip
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-262/Product: apolipoprotein A-I #status predicted <MAT>

Query Match 66.0%; Score 898.5; DB 2; Length 262;
Best Local Similarity 67.1%; Pred. No. 2.1e-43;
Matches 173; Conservative 37; Mismatches 45; Indels 3; Gaps 2;

QY 1 MKAAVLTAVLFLTGSQARHFQWDEPPQSPMDRVKDLATVYVDVLDKSGRDYVSQFEGS 60
Db 1 MKAAVLAVALVFLTGSQAHVWQODE-POSQMDKVDFANVYVDVAVKDSGRDYVSQFESS 59

QY 61 ALGKQLNLKLLDNWDSVTSTFSKLRQLGQVPTQEFWDNLEKETEGRLQKQMSKOLEEVKAK 120
Db 60 SLGQQLNLNLLENWDTLGSVLSQQLERLGLTRDFWNLLEKETDWDVRQEMNKOLEEVKQK 119

QY 121 VQPYLDDFOKKWQOEMELYRQKVEPLRAELQEGARQKLHELOKLSPLGEMDRDRAHV 180
Db 120 VQPYLDFQKKWQOEMELYRQKVEPLRAELQESARQKLQELQRLSPVAEEFDRMRTHV 179

QY 181 DALRTHLAPYSDELORLARLEALKENGARLAELYHAKATEHLSTLSEKAPALEDLRQ 240
Db 180 DSLRTQLAPHSEQMRESLAORLAELKSN--PTLNEVHTRAKTHLTKLGEKARPALEDLRH 237

QY 241 GLLPVLESFKVSFLSALE 258
Db 238 SLMPMLETLKTKAQSVIE 255

RESULT 15
A24700
apolipoprotein A-I precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 30-Jun-1988 #sequence revision 30-Jun-1988 #text_change 09-Jul-2004
C:Accession: A24700; S00298; A05314
R:Haddad, I.A.; Ordovas, J.M.; Fitzpatrick, T.; Karathanasis, S.K.
J. Biol. Chem. 261, 13268-13277, 1986
A:Title: Linkage, evolution, and expression of the rat apolipoprotein A-I, C-III, and A
A:Reference number: A92558; MUID:87008540; PMID:3020028
A:Accession: A24700
A:Molecule type: DNA
A:Residues: 1-259 <HAD>
A:Cross-references: UNIPROT:P04639; EMBL:J02597; NID:g202935; PIDN:AAA40745.1; PID:g202
R:Poncin, J.B.; Martial, J.A.; Gielen, J.E.
Eur. J. Biochem. 140, 493-498, 1984
A:Title: Cloning and structure analysis of the rat apolipoprotein A-I cDNA.
A:Reference number: S00298; MUID:84207987; PMID:6426556
A:Accession: S00298
A:Molecule type: mRNA
A:Residues: 1-259 <PON>

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2004, 07:13:28 ; Search time 157.462 Seconds
(without alignments)
975.635 Million cell updates/sec

Title: US-09-803-918A-2
Perfect score: 1362
Sequence: 1 MKAAVLTLAVLFLTGSQARH.....SPKVSFLSALEYTKKLNTQ 267

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot_02:.*
1: uniprot_sprot:.*
2: uniprot_trembl:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1362	100.0	267	1 APAL_HUMAN	P02647 homo sapien
2	1362	100.0	267	2 AAQ91811	Aa91811 homo sapi
3	1362	100.0	267	1 AAS68227	Aa68227 homo sapi
4	1306	95.9	267	1 APAL_MACFA	P15568 macaca fasc
5	1281	94.1	249	2 Q6LDN9	Q6ldn9 homo sapien
6	1281	94.1	249	2 AAAS1747	Aaas1747 homo sapi
7	1161.5	85.3	266	1 APAL_CANFA	P02648 canis famil
8	1105	81.1	265	1 APAL_PIG	P18648 sus scrofa
9	1100.5	80.8	266	1 APAL_RABIT	P09809 oryctolagus
10	1077	79.1	265	1 APAL_BOVIN	P15497 bos taurus
11	1012.5	74.3	284	2 Q9Z2L4	Q9z2l4 mesocricetu
12	975.5	71.6	285	1 APAL_TUPGB	O18759 tupaia glis
13	962	70.6	191	2 Q8HZ96	Q8hz96 gorilla gor
14	962	70.6	191	2 Q8HZ97	Q8hz97 pan troglod
15	947	69.5	191	2 Q8HZ95	Q8hz95 pongo pygma
16	905.5	66.4	264	2 Q8BPD5	Q8bpd5 mus muscullu
17	904.5	66.4	264	1 APAL_MOUSE	Q00623 mus muscullu
18	896	65.8	263	2 O08855	O08855 mus muscullu
19	895	65.7	263	2 O09042	O09042 mus muscullu
20	842.5	61.9	259	1 APAL_RAT	P04639 rattus norv
21	836	61.4	258	2 O09054	O09054 rattus norv
22	834	61.2	258	2 O08877	O08877 rattus norv
23	823	60.4	191	2 Q8HZ94	Q8hz94 saguinus oe
24	727.5	53.4	241	2 Q9TS49	Q9ts49 erinaceus e
25	705.5	51.8	264	1 APAL_ANAPL	O42296 anas platyr
26	688.5	50.6	264	1 APAL_CHICK	P08250 gallus gall
27	663.5	48.7	264	1 APAL_COTJA	P12918 coturnix co
28	526	38.6	260	2 Q7ZYS5	Q7zy5 xenopus lae
29	519	38.1	260	2 Q7S2A1	Q7s2a1 xenopus lae
30	512	37.6	261	2 Q6DDC5	Q6ddc5 xenopus tro
31	327	24.0	67	2 Q9Y355	Q9y355 homo sapien

32	320	23.5	79	2	Q6LD50
33	320	23.5	79	2	AAB35539
34	312.5	22.9	82	2	Q29248
35	306.5	22.5	367	2	Q6P7H6
36	306.5	22.5	367	2	AAB61667
37	303.5	22.3	250	2	Q6PBB8
38	303.5	22.3	250	2	AAB59786
39	298.5	21.9	263	2	Q98TG6
40	291	21.4	429	1	APA4_MACFA
41	279	20.5	260	2	Q6DHE2
42	275	20.2	262	1	APAL_BRARE
43	274.5	20.2	396	1	APA4_HUMAN
44	274.5	20.2	396	1	AAQ91809
45	274.5	20.2	396	2	AAS68228

ALIGNMENTS

RESULT 1
APAL_HUMAN
ID APAL_HUMAN STANDARD; PRT; 267 AA.
AC P02647;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Apolipoprotein A-I precursor (Apo-AI).
GN Name=APOA1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84221405; PubMed=6328445;
RA Sharpe C.R., Sidoli A., Shelley C.S., Lucero M.A., Shoulders C.C., Baralle F.E.;
RT "Human apolipoproteins AI, AII, CII and CIII. cDNA sequences and mRNA abundance.";
RL Nucleic Acids Res. 12:3917-3932(1984).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=85026665; PubMed=6207999;
RA Seilhamer J.J., Protter A.A., Frossard P., Levy-Wilson B.;
RT "Isolation and DNA sequence of full-length cDNA and of the entire gene for human apolipoprotein AI -- discovery of a new genetic polymorphism in the apo AI gene.";
RL DNA 3:309-317(1984).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=83220822; PubMed=6406984;
RA Shoulders C.C., Kornblitt A.R., Munro B.S., Baralle F.E.;
RT "Gene structure of human apolipoprotein AI.";
RL Nucleic Acids Res. 11:2827-2837(1983).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=83220772; PubMed=6304641;
RA Cheung P., Chan L.;
RT "Nucleotide sequence of cloned cDNA of human apolipoprotein A-I.";
RL Nucleic Acids Res. 11:3703-3715(1983).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=84119464; PubMed=6198645;
RA Law S.W., Brewer H.B. Jr.;
RT "Nucleotide sequence and the encoded amino acids of human apolipoprotein A-I mRNA.";
RL Proc. Natl. Acad. Sci. U.S.A. 81:66-70(1984).
RN [6]
RP SEQUENCE FROM N.A.
RX MEDLINE=86008382; PubMed=2995392;
RA Law S.W., Brewer H.B. Jr.;
RT "Tangier disease. The complete mRNA sequence encoding for preproapo-A-I.";

RL J. Biol. Chem. 260:12810-12814(1985).
 RN [7]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=84016011; PubMed=6413973;
 RA Karathanasis S.K., Zannis V.I., Breslow J.L.;
 RT "Isolation and characterization of the human apolipoprotein A-I
 gene.";
 RL Proc. Natl. Acad. Sci. U.S.A. 80:6147-6151(1983).
 RN [8]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89377481; PubMed=2673706;
 RA Moguilevsky N., Roobol C., Loriau R., Guillaume J.P., Jacobs P.,
 RA Cravador A., Herzog A., Browsers L., Scarso A., Gilles P.,
 RA Holmquist L., Carlson L.A., Bollen A.;
 RT "Production of human recombinant proapolipoprotein A-I in Escherichia
 coli: purification and biochemical characterization.";
 RL DNA 8:429-436(1989).
 RN [9]
 RP SEQUENCE FROM N.A. (VARIANT TANGIER).
 RX MEDLINE=88196137; PubMed=3129297;
 RA Makrides S.C., Ruiz-Opazo N., Hayden M., Nussbaum A.L., Breslow J.L.,
 RA Zannis V.I.;
 RT "Sequence and expression of Tangier apoA-I gene.";
 RL Eur. J. Biochem. 173:465-471(1988).
 RN [10]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Skeletal muscle;
 RX MEDLINE=22398257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.P., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Munz D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [11]
 RP SEQUENCE OF 118-267 FROM N.A.
 RX MEDLINE=83091059; PubMed=6294659;
 RA Breslow J.L., Ross D., McPherson J., Williams H.W., Kurnit D.,
 RA Nussbaum A.L., Karathanasis S.K., Zannis V.I.;
 RT "Isolation and characterization of cDNA clones for human
 apolipoprotein A-I.";
 RL Proc. Natl. Acad. Sci. U.S.A. 79:6861-6865(1982).
 RN [12]
 RP SEQUENCE OF 19-27.
 RX MEDLINE=83256553; PubMed=6409108;
 RA Brewer H.B. Jr., Fairwell T., Kay L., Meng M., Ronan R., Law S.,
 RA Light J.A.;
 RT "Human plasma proapoA-I: isolation and amino-terminal sequence.";
 RL Biochem. Biophys. Res. Commun. 113:626-632(1983).
 RN [13]
 RP SEQUENCE OF 25-267.
 RX MEDLINE=78123731; PubMed=204308;
 RA Brewer H.B. Jr., Fairwell T., Larue A., Ronan R., Houser A.,
 RA Bronzert T.J.;
 RT "The amino acid sequence of human APOA-I, an apolipoprotein isolated
 from high density lipoproteins.";
 RL Biochem. Biophys. Res. Commun. 80:623-630(1978).
 RN [14]
 RP SEQUENCE OF 25-267.
 RX MEDLINE=75133493; PubMed=164450;
 RA Baker H.N., Gotto A.M. Jr., Jackson R.L.;
 RT "The primary structure of human plasma high density apolipoprotein
 glutamine I (ApoA-I). II. The amino acid sequence and alignment of
 cyanogen bromide fragments IV, III, and I.";
 RL J. Biol. Chem. 250:2725-2738(1975).
 RN [15]
 RP SEQUENCE OF 25-56.
 RX MEDLINE=88331387; PubMed=3047170;
 RA Yui Y., Aoyama T., Morishita H., Takahashi M., Takatsu Y., Kawai C.;
 RT "Serum prostacyclin stabilizing factor is identical to apolipoprotein
 A-I (Apo A-I). A novel function of Apo A-I.";
 RL J. Clin. Invest. 82:803-807(1988).
 RN [16]
 RP SEQUENCE OF 25-48.
 RX MEDLINE=89380318; PubMed=2506184;
 RA Manjunath P., Marcel Y.L., Uma J., Seidah N.G., Chretien M.,
 RA Chapdelaine A.;
 RT "Apolipoprotein A-I binds to a family of bovine seminal plasma
 proteins.";
 RL J. Biol. Chem. 264:16853-16857(1989).
 RN [17]
 RP SEQUENCE OF 25-43.
 RX MEDLINE=88070603; PubMed=3120314;
 RA Prioli R.P., Ordovas J.M., Rosenberg I., Schaeffer E.J.,
 RA Pereira M.E.A.;
 RT "Similarity of cruzin, an inhibitor of Trypanosoma cruzi
 neuraminidase, to high-density lipoprotein.";
 RL Science 238:1417-1419(1987).
 RN [18]
 RP SEQUENCE OF 25-42.
 RC TISSUE=Heart;
 RX MEDLINE=95203287; PubMed=7895732;
 RA Corbett J.M., Wheeler C.H., Baker C.S., Yacoub M.H., Dunn M.J.;
 RT "The human myocardial two-dimensional gel protein database: update
 1994.";
 RL J. Biol. Chem. 261:3911-3914(1986).
 RN [20]
 RP PROCESSING.
 RX MEDLINE=83195100; PubMed=6405383;
 RA Zannis V.I., Karathanasis S.K., Keutmann H.T., Goldberger G.,
 RA Breslow J.L.;
 RT "Intracellular and extracellular processing of human apolipoprotein A-
 I: secreted apolipoprotein A-I isoprotein 2 is a propeptide.";
 RL Proc. Natl. Acad. Sci. U.S.A. 80:2574-2578(1983).
 RN [21]
 RP STRUCTURE BY NMR OF 190-209.
 RX MEDLINE=96270776; PubMed=8664326;
 RA Wang G., Treleaven W.D., Cushley R.J.;
 RT "Conformation of human serum apolipoprotein A-I(166-185) in the
 presence of sodium dodecyl sulfate or dodecylphosphocholine by ¹H-NMR
 and CD. Evidence for specific peptide-SDS interactions.";
 RL Biochim. Biophys. Acta 1301:174-184(1996).
 RN [22]
 RP X-RAY CRYSTALLOGRAPHY (4.0 ÅNGSTROMS) OF 67-267.
 RX MEDLINE=98024124; PubMed=9356442;
 RA Borhani D.W., Rogers D.P., Engler J.A., Brouillette C.G.;
 RT "Crystal structure of truncated human apolipoprotein A-I suggests a
 lipid-bound conformation.";
 RL Proc. Natl. Acad. Sci. U.S.A. 94:12291-12296(1997).
 RN [23]
 RP DISBASE.
 RX MEDLINE=94059128; PubMed=8240372;
 RA Nakata K., Kobayashi K., Yanagi H., Shimakura Y., Tsuchiya S.,
 RA Arinami T., Hamaguchi H.;
 RT "Autosomal dominant hypoalphalipoproteinemia due to a completely

RT defective apolipoprotein A-I gene."
 RL Biochem. Biophys. Res. Commun. 196:950-955 (1993).

Query Match 100.0%; Score 1362; DB 1; Length 267;
 Best Local Similarity 100.0%; Pred. No. 4.3e-67;
 Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKAAVLTAVLFLTGSOARHFQWQDEPPQSPWDRVKDLATVYVDVLKDSGRDYVSQFEPS 60
 DB 1 MKAAVLTAVLFLTGSOARHFQWQDEPPQSPWDRVKDLATVYVDVLKDSGRDYVSQFEPS 60

QY 61 ALGKQLNLKLLDNWDSVTSTFSKLRQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAK 120
 DB 61 ALGKQLNLKLLDNWDSVTSTFSKLRQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAK 120

QY 121 VQPYLDDFKKQWQEMELYRQKVEPLRAELQEGARQKLHELOEKLSPLGEMDRARAHV 180
 DB 121 VQPYLDDFKKQWQEMELYRQKVEPLRAELQEGARQKLHELOEKLSPLGEMDRARAHV 180

QY 181 DALRTHLAPYSDELRLQRLAARLEALKENGARLAAYHAKATEHLSTLSEKAKPALEDLRQ 240
 DB 181 DALRTHLAPYSDELRLQRLAARLEALKENGARLAAYHAKATEHLSTLSEKAKPALEDLRQ 240

QY 241 GLLPVLESFKVSFLSALEYTKKLTNQ 267
 DB 241 GLLPVLESFKVSFLSALEYTKKLTNQ 267

RESULT 2

AAQ91811 PRELIMINARY; PRT; 267 AA.

ID AAQ91811
 AC AAQ91811;
 DT 02-MAR-2004 (T-EMBLrel. 27, Created)
 DT 02-MAR-2004 (T-EMBLrel. 27, Last sequence update)
 DT 02-MAR-2004 (T-EMBLrel. 27, Last annotation update)
 DE Apolipoprotein A-I.
 GN APOA1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primata; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Nickerson D.A., Smith J.D., Fullerton S.M., Clark A.G., Stengard J.H.,
 RA Salomaa V., Boerwinkle E., Sing C.F., Weiss K.M.;
 RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY422952; AAQ91811.1; -
 KW Lipoprotein.
 SQ SEQUENCE 267 AA; 30778 MW; 1A28B8366E620310 CRC64;

Query Match 100.0%; Score 1362; DB 2; Length 267;
 Best Local Similarity 100.0%; Pred. No. 4.3e-67;
 Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKAAVLTAVLFLTGSOARHFQWQDEPPQSPWDRVKDLATVYVDVLKDSGRDYVSQFEPS 60
 DB 1 MKAAVLTAVLFLTGSOARHFQWQDEPPQSPWDRVKDLATVYVDVLKDSGRDYVSQFEPS 60

QY 61 ALGKQLNLKLLDNWDSVTSTFSKLRQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAK 120
 DB 61 ALGKQLNLKLLDNWDSVTSTFSKLRQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAK 120

QY 121 VQPYLDDFKKQWQEMELYRQKVEPLRAELQEGARQKLHELOEKLSPLGEMDRARAHV 180
 DB 121 VQPYLDDFKKQWQEMELYRQKVEPLRAELQEGARQKLHELOEKLSPLGEMDRARAHV 180

QY 181 DALRTHLAPYSDELRLQRLAARLEALKENGARLAAYHAKATEHLSTLSEKAKPALEDLRQ 240
 DB 181 DALRTHLAPYSDELRLQRLAARLEALKENGARLAAYHAKATEHLSTLSEKAKPALEDLRQ 240

QY 241 GLLPVLESFKVSFLSALEYTKKLTNQ 267
 DB 241 GLLPVLESFKVSFLSALEYTKKLTNQ 267

RESULT 3

AAS68227 PRELIMINARY; PRT; 267 AA.

ID AAS68227
 AC AAS68227;
 DT 29-MAR-2004 (T-EMBLrel. 27, Created)
 DT 29-MAR-2004 (T-EMBLrel. 27, Last sequence update)
 DT 29-MAR-2004 (T-EMBLrel. 27, Last annotation update)
 DE Apolipoprotein A-I.
 GN APOA1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Nickerson D.A., Smith J.D., Fullerton S.M., Buchanan A.V.,
 RA Sonpar V.A., Taylor S.L., Carlson C.S., Salomaa V., Stengard J.H.,
 RA Boerwinkle E., Clark A.G., Weiss K.M.;
 RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY555191; AAS68227.1; -
 KW Lipoprotein.
 SQ SEQUENCE 267 AA; 30778 MW; 1A28B8366E620310 CRC64;

Query Match 100.0%; Score 1362; DB 2; Length 267;
 Best Local Similarity 100.0%; Pred. No. 4.3e-67;
 Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKAAVLTAVLFLTGSOARHFQWQDEPPQSPWDRVKDLATVYVDVLKDSGRDYVSQFEPS 60
 DB 1 MKAAVLTAVLFLTGSOARHFQWQDEPPQSPWDRVKDLATVYVDVLKDSGRDYVSQFEPS 60

QY 61 ALGKQLNLKLLDNWDSVTSTFSKLRQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAK 120
 DB 61 ALGKQLNLKLLDNWDSVTSTFSKLRQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAK 120

QY 121 VQPYLDDFKKQWQEMELYRQKVEPLRAELQEGARQKLHELOEKLSPLGEMDRARAHV 180
 DB 121 VQPYLDDFKKQWQEMELYRQKVEPLRAELQEGARQKLHELOEKLSPLGEMDRARAHV 180

QY 181 DALRTHLAPYSDELRLQRLAARLEALKENGARLAAYHAKATEHLSTLSEKAKPALEDLRQ 240
 DB 181 DALRTHLAPYSDELRLQRLAARLEALKENGARLAAYHAKATEHLSTLSEKAKPALEDLRQ 240

QY 241 GLLPVLESFKVSFLSALEYTKKLTNQ 267
 DB 241 GLLPVLESFKVSFLSALEYTKKLTNQ 267

RESULT 4

APAL_MACFA STANDARD; PRT; 267 AA.

ID APAL_MACFA
 AC P15568; P17929;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Apolipoprotein A-I precursor (Apo-AI).
 GN Names=APOA1;
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey), and
 OS Papio hamadryas (Hamadryas baboon).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecinae; Macaca.
 OX NCBI_TaxID=9541, 9557;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=M.fascicularis;
 RX MEDLINE=87191989; PubMed=3106152;
 RA Polites H.G., Melchior G.W., Castle C.K., Marotti K.R.;
 RT "The primary structure of cynomolgus monkey apolipoprotein A-1 deduced
 from the cDNA sequence: comparison to the human sequence.";
 RL Gene 49:103-110 (1986).


```

Db      61  STFSKLEQPGVPTQEFWDNLEKTEGLRQEMSKOLEEVKAKVQPYLDLDDPQKKWQEMEL 120
QY      139 YRKQVEPLRAELOEGARQKLHELOEKLSPGCEEMDRARAHVDALRTHLAPYSDLRQRL 198
Db      121 YRKQVEPLRAELOEGARQKLHELOEKLSPGCEEMDRARAHVDALRTHLAPYSDLRQRL 180
QY      199 AARLEALKENGARLAAYHAKATEHLSTLSEKAKPALEDLRQGLLPVLESFKVSFSALE 258
Db      181 AARLEALKENGARLAAYHAKATEHLSTLSEKAKPALEDLRQGLLPVLESFKVSFSALE 240
QY      259 EYTKKLNTQ 267
Db      241 EYTKKLNTQ 249

RESULT 6
ID AAA51747 PRELIMINARY; PRT; 249 AA.
AC AAA51747;
DT 02-MAR-2004 (T-EMBLrel. 27, Created)
DT 02-MAR-2004 (T-EMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (T-EMBLrel. 27, Last annotation update)
DE APOAL protein (Fragment).
GN APOAL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Mogilevsky N., Roobol C., Loriau R., Guillaume J.P., Jacobs P.,
RA Cravador A., Herzog A., Brouwers L., Scarso A., Gilles P.,
RA Holmquist L., Carlson L.A., Bollen A.,
RL Submitted (APR-1990) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89377481; PubMed=2673706;
RA Mogilevsky N., Roobol C., Loriau R., Guillaume J.P., Jacobs P.,
RA Cravador A., Herzog A., Brouwers L., Scarso A., Gilles P.,
RA Holmquist L., Carlson L.A., Bollen A.,
RT "Production of human recombinant proapolipoprotein A-I in Escherichia
RT coli: purification and biochemical characterization."
RL DNA 8:429-436(1989).
DR EMBL; M29068; AAA51747.1; -.
FT NON TER 1 1
FT CHAIN 7 249 POTENTIAL.
SQ SEQUENCE 249 AA; 28961 MW; COA0B0B53903FAB4 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 1.1e-62;
Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      19  RHFWQDEPPQSPWDRVKDLATYVDVLKDSGRDYSVQFEGSALGKQLNKLKLDNDWSDVT 78
Db      1  RHFWQDEPPQSPWDRVKDLATYVDVLKDSGRDYSVQFEGSALGKQLNKLKLDNDWSDVT 60
QY      79  STFSKLEQPGVPTQEFWDNLEKTEGLRQEMSKOLEEVKAKVQPYLDLDDPQKKWQEMEL 138
Db      61  STFSKLEQPGVPTQEFWDNLEKTEGLRQEMSKOLEEVKAKVQPYLDLDDPQKKWQEMEL 120
QY      139 YRKQVEPLRAELOEGARQKLHELOEKLSPGCEEMDRARAHVDALRTHLAPYSDLRQRL 198
Db      121 YRKQVEPLRAELOEGARQKLHELOEKLSPGCEEMDRARAHVDALRTHLAPYSDLRQRL 180
QY      199 AARLEALKENGARLAAYHAKATEHLSTLSEKAKPALEDLRQGLLPVLESFKVSFSALE 258
Db      181 AARLEALKENGARLAAYHAKATEHLSTLSEKAKPALEDLRQGLLPVLESFKVSFSALE 240
QY      259 EYTKKLNTQ 267
Db      241 EYTKKLNTQ 249

RESULT 7
ID APAL CANFA STANDARD; PRT; 266 AA.
AC P02648;
DT 21-JUL-1986 (Rel. 01, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Apolipoprotein A-I precursor (Apo-AI).
GN Name=APOA1;
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OC NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=90332271; PubMed=2515239;
RA Luo C.-C., Li W.-H., Chan L.;
RT "Structure and expression of dog apolipoprotein A-I, E, and C-I mRNAs:
RT implications for the evolution and functional constraints of
RT apolipoprotein structure."
RL J. Lipid Res. 30:1735-1746(1989).
RN [2]
RP SEQUENCE OF 25-266.
RX MEDLINE=82142425; PubMed=6801039;
RA Chung H., Randolph A., Reardon I., Heinrikson R.L.;
RT "The covalent structure of apolipoprotein A-I from canine high density
RT lipoproteins."
RL J. Biol. Chem. 257:2961-2967(1982).
RN [3]
RP SEQUENCE OF 25-57 AND 262-265.
RX MEDLINE=76210910; PubMed=179887;
RA Nakai T., Whayne T.F., Tang J.;
RT "The amino- and carboxyl-terminal sequences of canine apolipoprotein
RT A-I."
RL FEBS Lett. 64:409-411(1976).
RN [4]
RP SEQUENCE OF 25-37.
RC TISSUE=Heart;
RX MEDLINE=98163340; PubMed=9504812;
RA Dunn M.J., Corbett J.M., Wheeler C.H.;
RT "HSC-2DPAGE and the two-dimensional gel electrophoresis database of
RT dog heart proteins."
RL Electrophoresis 18:2795-2802(1997).
CC -!- FUNCTION: Participates in the reverse transport of cholesterol
CC from tissues to the liver for excretion by promoting cholesterol
CC efflux from tissues and by acting as a cofactor for the lecithin
CC cholesterol acyltransferase (LCAT).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Major protein of plasma HDL, also found in
CC chylomicrons. Synthesized in the liver and small intestine.
CC -!- SIMILARITY: Belongs to the apolipoprotein AI/A4/E family.
DR PIR; A60940; LPDGA1.
DR HSP; P02647; IAVI.
DR HSC-2DPAGE; P02648; DOG.
DR InterPro; IPR000074; Apolipoprotein.
DR InterPro; IPR009074; Apolipo_A_E_C3.
DR Pfam; PF01442; Apolipoprotein; 1.
KW Cholesterol metabolism; Direct protein sequencing; HDL;
KW Lipid transport; Plasma; Repeat; Signal.
FT SIGNAL 1 18
FT PROPEP 19 24 By similarity.
FT CHAIN 25 266 Apolipoprotein A-I.
FT DOMAIN 67 266 10 X approximate tandem repeats.
FT REPEAT 67 88 1.
FT REPEAT 89 110 2.
FT REPEAT 111 121 3 (half-length).
FT REPEAT 122 143 4.
FT REPEAT 144 165 5.
FT REPEAT 166 187 6.
FT REPEAT 188 209 7.
FT REPEAT 210 231 8.

```

```
FT REPEAT      232 242      9 (half-length) .
FT REPEAT      243 266      10.
FT CONFLICT    168 168      E -> G (in Ref. 2) .
FT CONFLICT    202 202      E -> Q (in Ref. 2) .
FT CONFLICT    235 235      E -> Q (in Ref. 2) .
FT CONFLICT    264 266      NAQ -> A (in Ref. 3) .
SQ SEQUENCE    266 AA; A302620C28A869D CRC64;

Query Match      85.3%; Score 1161.5; DB 1; Length 266;
Best Local Similarity 85.0%; Pred. No. 4e-56;
Matches 227; Conservative 19; Mismatches 20; Indels 1; Gaps 1;

QY 1 MKAALVTLAVLFTGSGARHFQWQDRPPQSPWDRVKDLATVYVDVLKDSGRDVVSQFEGS 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MKAALVTLAVLFTGSGARHFQWQDRPPQSPWDRVKDLATVYVDVLKDSGRDVVSQFEGS 59
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 61 ALGKQLNLKLLDNWDSVTSTFSKLREQLGQVPTQEFWNLKETEGLRQENSKDLEEVKAK 120
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 60 ALGKQLNLKLLDNWDSVTSVTTKLRQIGVPTQEFWNLKETEGLRQENSKDLEEVKQK 119
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 121 VQPYLDDFQKKQWQEMELRYQKVEPLRAELQEGAROKLHLOQLKPLGEMMDRARAHV 180
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 120 VQPYLDDFQKKQWQEMELRYQKVEPLRAELQEGAROKLHLOQLKPLGEMMDRARAHV 179
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 181 DALRTHLAPYSDELRLAARLEALKENGARLAAYHAKATEHLSTLSKAKPALEDLRQ 240
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 180 DALRAQLAPYSDELRLAARLEALKENGARLAAYHAKATEHLSTLSKAKPALEDLRQ 239
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 241 GLLPVLESFKVSLSALEYTKKLNTQ 267
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 240 GLLPVLESFKVSLAAIDEATKLNLAQ 266
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 8
APAI_PIG
ID APAI_PIG STANDARD; PRT; 265 AA.
AC P186748;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Apolipoprotein A-I precursor (Apo-AI) .
GN Name=APOA1;
OS Sus scrofa (Pig) .
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
[1]
RN
RP MEDLINE=93224154; PubMed=8468059;
RX MEDLINE=94125128; PubMed=8294940;
RA Birchbauer A., Knipping G., Juritsch B., Aschauer H., Zechner R.;
RT "Characterization of the apolipoprotein AI and CIII genes in the
   domestic pig.";
RL Genomics 15:643-652(1993) .
[2]
RN
RP SEQUENCE FROM N.A.
RX MEDLINE=90132667; PubMed=2105375;
RA Weiler-Guettler H., Sommerfeldt M., Papandriopoulou A., Mischek U.,
RA Bonitz D., Frey A., Grupe M., Scheerer J., Gassen H.G.;
RT "Synthesis of apolipoprotein A-I in pig brain microvascular
   endothelial cells.";
RL J. Neurochem. 54:444-450(1990) .
[4]
RN
RP SEQUENCE OF 105-265 FROM N.A.
```

```
RC TISSUE=Liver;
RX MEDLINE=93154581; PubMed=8428656;
RT Trieu V.N., Hasler-Rapacz J., Rapacz J., Black D.D.;
RA "Sequences and expression of the porcine apolipoprotein A-I and C-III
   mRNAs.";
RL Gene 123:173-179(1993) .
[5]
RN
RP SEQUENCE OF 25-34.
RA Hasler-Rapacz J.O., Chaudhary R., Chowdhary B.P., Trieu V.N.,
RA Jackson K., Gustavsson I., Rapacz J.;
RL Submitted (OCT-1995) to Swiss-Prot.
[6]
RN
RP SEQUENCE OF 25-34.
RX MEDLINE=76184721; PubMed=178359;
RA Mahley R.W., Weisgraber K.H., Innerarity T., Brewer H.B. Jr.;
RT "Characterization of the plasma lipoproteins and apoproteins of the
   Erythrocebus patas monkey.";
RL Biochemistry 15:1928-1933(1976) .
CC -!- FUNCTION: Participates in the reverse transport of cholesterol
   from tissues to the liver for excretion by promoting cholesterol
   efflux from tissues and by acting as a cofactor for the lecithin
   cholesterol acyltransferase (LCAT) .
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Major protein of plasma HDL, also found in
   chylomicrons. Synthesized predominantly in the intestine and the
   liver.
CC -!- SIMILARITY: Belongs to the apolipoprotein AI/A4/E family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
   between the Swiss Institute of Bioinformatics and the EMBL outstation -
   the European Bioinformatics Institute. There are no restrictions on its
   use by non-profit institutions as long as its content is in no way
   modified and this statement is not removed. Usage by and for commercial
   entities requires a license agreement (See http://www.isb-sib.ch/announce/
   or send an email to license@isb-sib.ch)
CC
CC EMBL; L00626; AAA30992.1; -
DR EMBL; X69477; CAA49234.1; -
DR EMBL; X17057; -; NOT ANNOTATED_CDS.
DR EMBL; X59414; CAA42050.1; -
DR PIR; A46018; A46018.
DR PIR; JT0672; JT0672.
DR PIR; S21830; S21830.
DR PIR; S31394; S31394.
DR HSSP; P02647; IAVI.
DR InterPro; IPR000074; Apolipoprotein.
DR InterPro; IPR009074; Apolipo_A_E_C3.
DR Pfam; PF01442; Apolipoprotein; 1.
KW Cholesterol metabolism; Direct protein sequencing; HDL;
KW Lipid transport; Plasma; Repeat; Signal.
FT SIGNAL 1 18 By similarity.
FT PROPEP 19 24
FT CHAIN 25 265
FT DOMAIN 67 265
FT REPEAT 67 88 1.
FT REPEAT 89 110 2.
FT REPEAT 111 121 3 (half-length) .
FT REPEAT 122 142 4.
FT REPEAT 144 165 5.
FT REPEAT 166 187 6.
FT REPEAT 188 209 7.
FT REPEAT 210 230 8.
FT REPEAT 231 241 9 (half-length) .
FT REPEAT 242 265 10.
FT CONFLICT 108 108 E -> K (in Ref. 3) .
FT CONFLICT 143 143 Missing (in Ref. 2 and 3) .
FT CONFLICT 173 173 D -> S (in Ref. 4) .
FT CONFLICT 180 180 E -> A (in Ref. 2 and 3) .
FT CONFLICT 185 186 HV -> QL (in Ref. 1 and 5) .
FT CONFLICT 209 209 G -> D (in Ref. 2 and 3) .
FT CONFLICT 224 224 A -> G (in Ref. 4) .
SQ SEQUENCE 265 AA; 2C6E578318ECF69C CRC64;
```

Query Match 81.1%; Score 1105; DB 1; Length 265;
Best Local Similarity 80.9%; Pred. No. 4.9e-53;
Matches 216; Conservative 20; Mismatches 29; Indels 2; Gaps 2;

QY 1 MKAAVLTAVLFLTGSQARHFWDSPQSPDRVKDLATVYVDVKDSGRDYVSQFEAS 60
DB 1 MKAAVLTAVLFLTGSQARHFWDSPQSPDRVKDLATVYVDVKDSGRDYVSQFEAS 59
QY 61 ALGKQLNLKLLDNWDSVTSTFSKLRQLGPVTOEFWDNLEKETEGRLQEMSKDLEEVKAK 120
DB 60 ALGKHLNLKLLDNWDSLGSFTFKVREQGPVTOEFWDNLEKETEGRLQEMSKDLEEVKAK 119
QY 121 VQPYLDDFQKKWQEMELVYRKVEPLRAELQEGARQKLHLEQKLSPLGEMDRARAHV 180
DB 120 VQPYLDDFQKKWQEMELVYRKVEPLRAELQEGARQKLHLEQKLSPLGEMDRARAHV 179
QY 181 DALRTHLAPYSDELQRLAARLEALKENGARLAHYHAKATEHLSTLSEKAKPALEDLRQ 240
DB 180 EALRQHVAPYSDDLORMAARFALKEGGGS-LAEYQAKAQEQKALGEKAKPALEDLRQ 238
QY 241 GLLPVLESFKVSLSALEEVYTKLNTQ 267
DB 239 GLLPVLENLKVSLIAAIDEASKKLNQ 265

RESULT 9

APAL_RABBIT STANDARD; PRT; 266 AA.

AC P09809; 01-MAR-1989 (Rel. 10, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Apolipoprotein A-I precursor (Apo-AI).
GN Name=APOA1;
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=22AP AI; TISSUE=Small intestine;
RA Paraskevopoulou T.B., Kritis A., Zannis V.I.;
RL Submitted (JUL-1989) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Intestine;
RX MEDLINE=88082866; PubMed=3121329;
RA Fan T.C., Hao Q.L., Yamin T.T., Dai P.H., Chen B.S., Chen S.L.,
RA Kroon P.A., Chao Y.S.;
RT "Rabbit apolipoprotein A-I mRNA and gene. Evidence that rabbit
RT apolipoprotein A-I is synthesized in the intestine but not in the
RT liver";
RL Eur. J. Biochem. 170:99-104 (1987).
RN [3]
RP SEQUENCE OF 25-266.
RX MEDLINE=87030294; PubMed=30951115;
RA Yang C., Yang T., Pownall H.J., Gotto A.M. Jr.;
RT "The primary structure of apolipoprotein A-I from rabbit high-density
RT lipoprotein";
RL Eur. J. Biochem. 160:427-431 (1986).
CC -!- FUNCTION: Participates in the reverse transport of cholesterol
CC efflux from tissues to the liver for excretion by promoting cholesterol
CC cholesterol acyltransferase (LCAT).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Major protein of plasma HDL, also found in
CC chylomicrons.
CC -!- SIMILARITY: Belongs to the apolipoprotein A1/A4/E family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL; X15908; CAA34024.1; -;
DR EMBL; X06658; CAA29857.1; -;
DR EMBL; X06659; CAA29858.1; -;
DR PIR; S06064; LPRB1Z.
DR HSSP; P02647; IAVI.
DR InterPro; IPR000074; Apolipoprotein.
DR InterPro; IPR009074; Apolipo_A_E_C3.
DR Pfam; PF01442; Apolipoprotein; 1.
KW Cholesterol metabolism; Direct protein sequencing; HDL;
KW Lipid transport; Plasma; Repeat; Signal.
FT SIGNAL 1 18
FT PROPEP 19 24
FT CHAIN 25 266 Apolipoprotein A-I.
FT DOMAIN 67 266 10 X approximate tandem repeats.
FT REPEAT 67 98 1.
FT REPEAT 89 110 2.
FT REPEAT 111 121 3 (half-length).
FT REPEAT 122 143 4.
FT REPEAT 144 165 5.
FT REPEAT 166 187 6.
FT REPEAT 188 209 7.
FT REPEAT 210 231 8.
FT REPEAT 232 242 9 (half-length).
FT REPEAT 243 266 10.
FT CONFLICT 18 18 A -> R (in Ref. 2; CAA29858).
FT CONFLICT 44 44 Missing (in Ref. 3).
FT CONFLICT 45 45 V -> I (in Ref. 2; CAA29858).
FT CONFLICT 107 107 E -> Q (in Ref. 3).
FT CONFLICT 123 123 Y -> F (in Ref. 2; CAA29857).
FT CONFLICT 147 147 A -> V (in Ref. 2; CAA29858 and 3).
FT CONFLICT 150 150 R -> G (in Ref. 2; CAA29858 and 3).
FT CONFLICT 191 191 N -> Q (in Ref. 3).
FT CONFLICT 195 195 Missing (in Ref. 2; CAA29857).
FT CONFLICT 211 211 S -> K (in Ref. 2; CAA29858 and 3).
FT CONFLICT 255 256 VL -> LV (in Ref. 3).
FT CONFLICT 256 256 L -> V (in Ref. 2; CAA29857).
SQ SEQUENCE 266 AA; 30591 MW; 0FF6DB386497C7D2 CRC64;
Query Match 80.8%; Score 1100.5; DB 1; Length 266;
Best Local Similarity 80.1%; Pred. No. 8.6e-53;
Matches 214; Conservative 25; Mismatches 27; Indels 1; Gaps 1;

QY 1 MKAAVLTAVLFLTGSQARHFWDSPQSPDRVKDLATVYVDVKDSGRDYVSQFEAS 60
DB 1 MKAAVLTAVLFLTGSQARHFWDSPQSPDRVKDLATVYVDVKDSGRDYVSQFEAS 59
QY 61 ALGKQLNLKLLDNWDSVTSTFSKLRQLGPVTOEFWDNLEKETEGRLQEMSKDLEEVKAK 120
DB 60 ALGKQLNLKLLDNWDSVTSTFSKLRQLGPVTOEFWDNLEKETEGRLQEMSKDLEEVKAK 119
QY 121 VQPYLDDFQKKWQEMELVYRKVEPLRAELQEGARQKLHLEQKLSPLGEMDRARAHV 180
DB 120 VQPYLDDFQKKWQEMELVYRKVEPLRAELQEGARQKLHLEQKLSPLGEMDRARAHV 179
QY 181 DALRTHLAPYSDELQRLAARLEALKENGARLAHYHAKATEHLSTLSEKAKPALEDLRQ 240
DB 180 DTLRKLAPYSNELQRLAARLESIKEGGASLAELYQAKAREHLSVLSEKAPALEDLRQ 239
QY 241 GLLPVLESFKVSLSALEEVYTKLNTQ 267
DB 240 GLLPVLESFKVSLSALEEVYTKLNTQ 266

RESULT 10
APAL_BOVIN STANDARD; PRT; 265 AA.
AC P15497;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)

05-JUL-2004 (Rel. 44, Last annotation update)
 Apolipoprotein A-I precursor (Apo-AI).
 Name=ApoAI;
 Bos taurus (Bovine).
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovinae; Bos.
 NCBI_TaxID=9913;
 (1)
 SEQUENCE FROM N.A.
 MEDLINE=90348478; PubMed=2117227;
 O'Huigin C., Chan L., Li W.H.;
 "Cloning and sequencing of bovine apolipoprotein A-I cDNA and
 molecular evolution of apolipoproteins A-I and B-100.";
 Mol. Biol. Evol. 7:327-339(1990).
 (2)
 SEQUENCE OF 19-265.
 MEDLINE=92153895; PubMed=179745;
 Sparrow D.A., Lee B.R., Laplaud M.P., Auboirion S., Bauchart D.,
 Chapman J.M., Gotto A.M. Jr., Yang C.Y., Sparrow J.T.;
 "Plasma lipid transport in the prerinant calf, Bos spp: primary
 structure of bovine apolipoprotein A-I.";
 Biochim. Biophys. Acta 1123:145-150(1992).
 (3)
 SEQUENCE OF 25-70.
 MEDLINE=90147795; PubMed=2105728;
 Auboirion S., Sparrow D.A., Beaubatie L., Bauchart D., Sparrow J.T.,
 Laplaud M.P., Chapman J.M.;
 "Characterization and amino-terminal sequence of apolipoprotein AI
 from plasma high density lipoproteins in the prerinant calf, Bos
 spp.";
 Biochem. Biophys. Res. Commun. 166:833-839(1990).
 -!- FUNCTION: Participates in the reverse transport of cholesterol
 from tissues to the liver for excretion by promoting cholesterol
 efflux from tissues and by acting as a cofactor for the lecithin
 cholesterol acyltransferase (LCAT).
 -!- SUBCELLULAR LOCATION: Secreted.
 -!- TISSUE SPECIFICITY: Major protein of plasma HDL, also found in
 chylomicrons.
 -!- SIMILARITY: Belongs to the apolipoprotein A1/A4/E family.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as the content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 or send an email to license@isb-sib.ch).

 EMBL; M35870; AAA30381.1; -;
 PIR; I45853; A56858.
 HSP; P02647; IAV1.
 InterPro; IPR000074; Apolipoprotein.
 InterPro; IPR009074; Apolipo A E C3.
 Pfam; PF01442; Apolipoprotein; 1.
 Cholesterol metabolism; Direct protein sequencing; HDL;
 Lipid transport; Plasma; Repeat; Signal.
 SIGNAL 1 18
 PROPEP 19 24 Apolipoprotein A-I.
 CHAIN 25 265 10 x approximate tandem repeats.
 DOMAIN 67 265 1.
 REPEAT 67 88 1.
 REPEAT 89 110 2.
 REPEAT 111 121 3 (half-length).
 REPEAT 122 143 4.
 REPEAT 144 165 5.
 REPEAT 166 187 6.
 REPEAT 188 209 7.
 REPEAT 210 230 8.
 REPEAT 231 241 9 (half-length).
 REPEAT 242 265 10.
 CONFLICT 185 186 QL -> HV (in Ref. 2).
 SEQUENCE 265 AA; 30276 MW; 06A2681EA2ABA50F CRC64;

Query Match 79.1%; Score 1077; DB 1; Length 265;
 Best Local Similarity 78.7%; Pred. No. 1.7e-51;
 Matches 210; Conservative 23; Mismatches 32; Indels 2; Gaps 2;
 QY 1 MKAAVLTAVLFLTGSQARHFHQDDPPQSPMDRVKDLATVYVDVVKOSGRDYVSQFEGS 60
 DB 1 MKAAVLTAVLFLTGSQARHFHQDDPPQSPMDRVKDLATVYVDVVKOSGRDYVSQFEGS 59
 QY 61 ALGKQINLKLNDWDSVTSTFSKLEQGLPVTVQEFWDNLEKTEGLRQMSKDLSEVKAK 120
 DB 60 ALGKQINLKLNDWDSVTSTFSKLEQGLPVTVQEFWDNLEKTEGLRQMSKDLSEVKAK 119
 QY 121 VQPYLDDFQKKWQEWEMELYRQKVEPLRAELQEGARQKLHELOEKLSPGLGEEMDRARAHV 180
 DB 120 VQPYLDDFQKKWQEWEMELYRQKVEPLRAELQEGARQKLHELOEKLSPGLGEEMDRARAHV 179
 QY 181 DALRTHLPYSDELQRLAARLEALKENGAGARLAHYHAKATEHLSTLSEKAKPALEDLRQ 240
 DB 180 ETLRQQLAPYSDDLQRLTARLEALKEGGS-LAEYHAKASQLKALGEKAKPVLEDLRQ 238
 QY 241 GLLPVLESFKVSFLSALSEYTKKLNQ 267
 DB 239 GLLPVLESFKVSFLSALSEYTKKLNQ 265
 RESULT 11
 Q9Z2L4 PRELIMINARY; PRT; 264 AA.
 AC Q9Z2L4; (TREMREL. 10, Created)
 DT 01-MAY-1999 (TREMREL. 10, Last sequence update)
 DT 01-MAY-1999 (TREMREL. 10, Last sequence update)
 DT 01-MAY-2004 (TREMREL. 26, Last annotation update)
 DE Apolipoprotein A-I.
 GN Name=ApoAI;
 OS Mesocricetus auratus (Golden hamster).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
 OC Mesocricetus.
 OC NCBI_TaxID=10036;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Golden Syrian; TISSUE=Intestine;
 RX MEDLINE=99061559; PubMed=9843713;
 RA Wu J.Y.J., Reeves S.K., Wang Y.R., Wu Y., Lei P.P., Lei K.Y.;
 "Zinc deficiency decreases plasma level and hepatic mRNA abundance of
 apolipoprotein A-I in rats and hamsters.";
 Am. J. Physiol. 275:0-0(1998).
 DR EMBL; AF046919; AAC98484.1; -;
 DR HSP; P02647; IAV1.
 DR GO; GO:0005576; C:extracellular; IEA.
 DR GO; GO:0008289; F:lipid binding; IEA.
 DR GO; GO:0006869; P:lipid transport; IEA.
 DR GO; GO:0042157; P:lipoprotein metabolism; IEA.
 DR InterPro; IPR000074; Apolipoprotein.
 DR InterPro; IPR009074; Apolipo A E C3.
 DR Pfam; PF01442; Apolipoprotein; 1.
 KW Lipoprotein.
 SQ SEQUENCE 264 AA; 30739 MW; 280B22F4C0F0B129 CRC64;
 Query Match 74.3%; Score 1012.5; DB 2; Length 264;
 Best Local Similarity 73.0%; Pred. No. 5.6e-48;
 Matches 195; Conservative 30; Mismatches 39; Indels 3; Gaps 2;
 QY 1 MKAAVLTAVLFLTGSQARHFHQDDPPQSPMDRVKDLATVYVDVVKOSGRDYVSQFEGS 60
 DB 1 MKAAVLTAVLFLTGSQARHFHQDDPPQSPMDRVKDLATVYVDVVKOSGRDYVSQFEGS 59
 QY 61 ALGKQINLKLNDWDSVTSTFSKLEQGLPVTVQEFWDNLEKTEGLRQMSKDLSEVKAK 120
 DB 60 ALGKQINLKLNDWDSVTSTFSKLEQGLPVTVQEFWDNLEKTEGLRQMSKDLSEVKAK 119
 QY 121 VQPYLDDFQKKWQEWEMELYRQKVEPLRAELQEGARQKLHELOEKLSPGLGEEMDRARAHV 180


```
Q8HZ97
ID Q8HZ97 PRELIMINARY; PRT; 191 AA.
AC Q8HZ97;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Apolipoprotein A-I (Fragment).
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RA O'huigin C., Tichy H., Klein J.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY092007; AAM76624.1; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0008289; F:lipid binding; IEA.
DR GO; GO:0006869; P:lipid transport; IEA.
DR GO; GO:0042157; P:lipoprotein metabolism; IEA.
DR InterPro; IPR000074; Apolipoprotein.
DR InterPro; IPR009074; Apolipo A_E C3.
DR Pfam; PF01442; Apolipoprotein; 1.
KW Lipoprotein.
FT NON_TER 1 1
FT SEQUENCE 191 AA; 22087 MW; 2D30919874C8F0F4 CRC64;
SQ SEQUENCE 191 AA; 22087 MW; 2D30919874C8F0F4 CRC64;

Query Match 70.6%; Score 962; DB 2; Length 191;
Best Local Similarity 99.0%; Pred. No. 2.3e-45;
Matches 189; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 38 LATVYVDVLKSGRDYVSQFEGSALGKQLNLKLLDNWDSVTSTFSKLRQLGPGVTQEFWD 97
DB 1 LATVYVDVLKSGRDYVSQFEGSALGKQLNLKLLDNWDSVTSTFSKLRQLGPGVTQEFWD 60

QY 98 NLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQEEEMELYRQKVEPLRAELQEGAROK 157
DB 61 NLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQEEEMELYRQKVEPLRAELQEGAROK 120

QY 158 LHELOEKLSPGGEEMDRARAHVDALRTHLAPYSDELQRQRLAARLEALKENGSGARLAIEYH 217
DB 121 LHELOEKLSPGGEEMDRARAHVDALRTHLAPYSDELQRQRLAARLEALKENGSGARLAIEYH 180

QY 218 AKATEHLSTLS 228
DB 181 AKATEHLSTLS 191

RESULT 15
Q8HZ95
ID Q8HZ95 PRELIMINARY; PRT; 191 AA.
AC Q8HZ95;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Apolipoprotein A-I (Fragment).
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP SEQUENCE FROM N.A.
RA O'huigin C., Tichy H., Klein J.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY092009; AAM76626.1; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0008289; F:lipid binding; IEA.
DR GO; GO:0006869; P:lipid transport; IEA.
DR GO; GO:0042157; P:lipoprotein metabolism; IEA.
DR InterPro; IPR000074; Apolipoprotein.
DR InterPro; IPR009074; Apolipo A_E C3.
DR Pfam; PF01442; Apolipoprotein; 1.
KW Lipoprotein.
FT NON_TER 1 1
FT SEQUENCE 191 AA; 22087 MW; 2D30919874C8F0F4 CRC64;
SQ SEQUENCE 191 AA; 22087 MW; 2D30919874C8F0F4 CRC64;

Query Match 69.5%; Score 947; DB 2; Length 191;
Best Local Similarity 96.9%; Pred. No. 1.5e-44;
Matches 185; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 38 LATVYVDVLKSGRDYVSQFEGSALGKQLNLKLLDNWDSVTSTFSKLRQLGPGVTQEFWD 97
DB 1 LATVYVDVLKSGRDYVSQFEGSALGKQLNLKLLDNWDSVTSTFSKLRQLGPGVTQEFWD 60

QY 98 NLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQEEEMELYRQKVEPLRAELQEGAROK 157
DB 61 NLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQEEEMELYRQKVEPLRAELQEGAROK 120

QY 158 LHELOEKLSPGGEEMDRARAHVDALRTHLAPYSDELQRQRLAARLEALKENGSGARLAIEYH 217
DB 121 LHELOEKLSPGGEEMDRARAHVDALRTHLAPYSDELQRQRLAARLEALKENGSGARLAIEYH 180

QY 218 AKATEHLSTLS 228
DB 181 AKATEHLSTLS 191

Search completed: December 21, 2004, 07:40:27
Job time : 159.462 secs
```

```
KW Lipoprotein.
FT NON_TER 1 1
FT SEQUENCE 191 AA; 22128 MW; DBD0671444C13725 CRC64;
SQ SEQUENCE 191 AA; 22128 MW; DBD0671444C13725 CRC64;

Query Match 69.5%; Score 947; DB 2; Length 191;
Best Local Similarity 96.9%; Pred. No. 1.5e-44;
Matches 185; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 38 LATVYVDVLKSGRDYVSQFEGSALGKQLNLKLLDNWDSVTSTFSKLRQLGPGVTQEFWD 97
DB 1 LATVYVDVLKSGRDYVSQFEGSALGKQLNLKLLDNWDSVTSTFSKLRQLGPGVTQEFWD 60

QY 98 NLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQEEEMELYRQKVEPLRAELQEGAROK 157
DB 61 NLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQEEEMELYRQKVEPLRAELQEGAROK 120

QY 158 LHELOEKLSPGGEEMDRARAHVDALRTHLAPYSDELQRQRLAARLEALKENGSGARLAIEYH 217
DB 121 LHELOEKLSPGGEEMDRARAHVDALRTHLAPYSDELQRQRLAARLEALKENGSGARLAIEYH 180

QY 218 AKATEHLSTLS 228
DB 181 AKATEHLSTLS 191

Search completed: December 21, 2004, 07:40:27
Job time : 159.462 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.
OM protein - protein search, using sw model
Run on: December 21, 2004, 06:24:01 ; Search time 159.744 Seconds
(without alignments)
599.590 Million cell updates/sec

Title: US-09-803-918A-2
Perfect score: 1362
Sequence: 1 MKAALVTLAVLTGSLQARH.....SPKVSFLSALEYTKYLTNQ 267

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:*

- 1: Genesecp1980s:*
- 2: Genesecp1990s:*
- 3: Genesecp2000s:*
- 4: Genesecp2001s:*
- 5: Genesecp2002s:*
- 6: Genesecp2003as:*
- 7: Genesecp2003bs:*
- 8: Genesecp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1362	100.0	267	1 AAP61079	Aap61079 Assumed h
2	1362	100.0	267	1 AAP82128	Aap82128 Entire hu
3	1362	100.0	267	2 AAP72705	Aar72705 Human apo
4	1362	100.0	267	2 AAP34032	Aar34032 Sequence
5	1362	100.0	267	2 AAY18675	Aay18675 Human apo
6	1362	100.0	267	4 AAB47620	Aab47620 Full leng
7	1362	100.0	267	5 AAO15892	Aao15892 Human apo
8	1362	100.0	267	5 ABG97593	Abg97593 Human apo
9	1362	100.0	267	6 ABR44031	Abra44031 Human apo
10	1362	100.0	267	6 ABP57065	Abp57065 Human apo
11	1362	100.0	267	6 ADA61214	Ada61214 Human apo
12	1362	100.0	267	7 ADJ83084	Adj83084 Human apo
13	1362	100.0	267	8 ADI19752	Adi19752 Human apo
14	1362	100.0	275	6 AAO30162	Aao30162 Human apo
15	1359	99.8	267	7 ADJ68447	Adj68447 Human hea
16	1359	99.8	267	7 ADJ83083	Adj83083 Human pro
17	1355	99.5	267	6 ABR44032	Abra44032 Human mut
18	1354	99.4	267	2 AAW08602	Aaw08602 Human apo
19	1351.5	99.2	268	1 AAP80668	Aap80668 Recombina
20	1346	98.8	267	8 ADE76862	Ade76862 Human pro
21	1328	97.5	299	4 AAU33170	Aau33170 Novel hum
22	1306	95.9	267	7 ADJ83086	Adj83086 Apolipop
23	1299	95.4	267	7 ADJ83087	Adj83087 Crab-eati
24	1281	94.1	249	7 ADJ83085	Adj83085 Human pro
25	1276	93.7	250	5 AAE24642	Aae24642 Human Pro

26	1268.5	93.1	252	6 ABR43302	Abra43302 Human lip
27	1246.5	91.5	306	5 ABG97587	Abg97587 Human apo
28	1246.5	91.5	325	5 ABG97603	Abg97603 Human apo
29	1246	91.5	304	5 ABG97586	Abg97586 Human apo
30	1246	91.5	304	5 ABG97585	Abg97585 Human apo
31	1246	91.5	323	5 ABG97602	Abg97602 Human apo
32	1246	91.5	323	5 ABG97601	Abg97601 Human apo
33	1245	91.4	254	7 ADD29962	Add29962 Mature hu
34	1243.5	91.3	304	5 ABG97584	Abg97584 Human apo
35	1243.5	91.3	306	5 ABG97588	Abg97588 Human apo
36	1243.5	91.3	306	5 ABG97589	Abg97589 Human apo
37	1243.5	91.3	323	5 ABG97600	Abg97600 Human apo
38	1243.5	91.3	325	5 ABG97604	Abg97604 Human apo
39	1243.5	91.3	325	5 ABG97605	Abg97605 Human apo
40	1242.5	91.2	337	5 ABG97595	Abg97595 Human apo
41	1242	91.2	301	5 ABG97583	Abg97583 Human apo
42	1242	91.2	301	5 ABG97581	Abg97581 Human apo
43	1242	91.2	316	5 ABG97599	Abg97599 Human apo
44	1242	91.2	316	5 ABG97596	Abg97596 Human apo
45	1242	91.2	329	5 ABG97592	Abg97592 Human apo

ALIGNMENTS

RESULT 1
AAP61079
ID AAP61079 standard; protein; 267 AA.
XX
XX AAP61079;
XX AC
XX 25-MAR-2003 (revised)
DT 07-OCT-1991 (first entry)
XX
XX
DE Assumed human apolipoprotein A-1 derivative gene product.
XX
XX Hyperlipaemia; arteriosclerosis.
XX
XX Homo sapiens.
XX
XX JP61096998-A.
XX
XX 15-MAY-1986.
XX
XX 16-OCT-1984; 84JP-00216988.
XX
XX 16-OCT-1984; 84JP-00216988.
XX
XX (MITU) MITSUBISHI CHEM IND LTD.
XX
XX WPI; 1986-165025/26.
XX N-PSDB; AAN60886.

Human apo:lipoprotein A-1 (deriv.) prepn. - by providing DNA fragment in cloning site downstream of expression vector promoter and introducing into host microorganism.
XX
XX Disclosure; Fig 2; 9pp; Japanese.
XX
XX The human apolipoprotein may be produced by a suitable transformed host, it is effective in treating hyperlipaemia and arteriosclerosis. (Updated on 25-MAR-2003 to correct PA field.)
CC
XX
XX
SQ Sequence 267 AA;

Query Match 100.0%; Score 1362; DB 1; Length 267;
Best Local Similarity 100.0%; Pred. No. 3e-99;
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MKAALVTLAVLTGSLQARHFWQDEPPQSPWDRVKDLATVYDVLKDSGRDYVSQEGS 60
Db 1 MKAALVTLAVLTGSLQARHFWQDEPPQSPWDRVKDLATVYDVLKDSGRDYVSQEGS 60

QY 61 ALGKQLNLKLDNWDVSTSTFSKLRQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAK 120
 |||||||
 Db 61 ALGKQLNLKLDNWDVSTSTFSKLRQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAK 120
 |||||||
 QY 121 VQPYLDDDFQKQWQEMELYRQKVEPLRAELQEGARQKLHELOEKLSPLGGEEMRDRARAHV 180
 |||||||
 Db 121 VQPYLDDDFQKQWQEMELYRQKVEPLRAELQEGARQKLHELOEKLSPLGGEEMRDRARAHV 180
 |||||||
 QY 181 DALRTHLAPYSDELQRRLAARLEALKENGAGARLAHYHAKATEHLSLSEKAKPALEDLRQ 240
 |||||||
 Db 181 DALRTHLAPYSDELQRRLAARLEALKENGAGARLAHYHAKATEHLSLSEKAKPALEDLRQ 240
 |||||||
 QY 241 GLLPVLESFKVSFLSALEEYTKKLNQ 267
 |||||||
 Db 241 GLLPVLESFKVSFLSALEEYTKKLNQ 267
 |||||||

RESULT 2
 AAP82128
 ID AAP82128 standard; protein; 267 AA.
 XX
 AC AAP82128;
 XX
 DT 25-MAR-2003 (revised)
 DT 24-OCT-1990 (first entry)
 XX
 XX Entire human preproapoprotein A1.
 DE
 XX human preproapoprotein A1; high density lipoprotein deficiency; ss.
 KW
 XX Synthetic.
 OS
 FH Key Location/Qualifiers
 FT Peptide 1..18
 FT /label= precursor
 FT Peptide 19..24
 FT /label= propeptide
 FT Protein 25..267
 FT /label= mature apoprotein
 FT
 XX EP293357-A.
 XX
 XX 30-NOV-1988.
 XX
 XX 24-MAY-1988; 88EP-00870095.
 XX
 XX 28-MAY-1987; 87GB-00012540.
 XX
 XX (UNIO) UCB SA.
 XX (BOLL/) BOLLEN A.
 XX
 XX Bollen A, Gobert J, Wulfert E;
 PI
 XX WPI; 1988-339891/48.
 DR
 XX N-PSDB; AAN82064.
 XX
 XX New DNA encoding human pro-apo-lipoprotein A1 - modified to eliminate
 PT hairpin structures.
 FT
 XX Disclosure; Page ?; 25pp; French.
 PS
 XX The cDNA 878bp fragment encoding preproapoprotein A1 was detected in
 CC clone pUBI609 derived from human liver cells. See also AAN81258.
 CC (Updated on 25-MAR-2003 to correct PA field.)
 CC
 XX Sequence 267 AA;
 SQ

Query Match 100.0%; Score 1362; DB 1; Length 267;
 Best Local Similarity 100.0%; Pred. No. 3e-99;
 Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKAAVLTIAVLFTGSGQARHFQQDEPPQSPMDRVKDLATVYVDLKDGRDYSQFEGS 60
 |||||||
 Db 61 ALGKQLNLKLDNWDVSTSTFSKLRQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAK 120
 |||||||
 QY 121 VQPYLDDDFQKQWQEMELYRQKVEPLRAELQEGARQKLHELOEKLSPLGGEEMRDRARAHV 180
 |||||||
 Db 121 VQPYLDDDFQKQWQEMELYRQKVEPLRAELQEGARQKLHELOEKLSPLGGEEMRDRARAHV 180
 |||||||
 QY 181 DALRTHLAPYSDELQRRLAARLEALKENGAGARLAHYHAKATEHLSLSEKAKPALEDLRQ 240
 |||||||
 Db 181 DALRTHLAPYSDELQRRLAARLEALKENGAGARLAHYHAKATEHLSLSEKAKPALEDLRQ 240
 |||||||
 QY 241 GLLPVLESFKVSFLSALEEYTKKLNQ 267
 |||||||
 Db 241 GLLPVLESFKVSFLSALEEYTKKLNQ 267
 |||||||

RESULT 3
 AAR72705
 ID AAR72705 standard; protein; 267 AA.
 XX
 AC AAR72705;
 XX
 DT 31-OCT-1995 (first entry)
 XX
 XX Human apo A-I including signal and propeptide sequences.
 DE
 XX Apo A-I; LDL cholesterol; low density lipoprotein; lipid.
 KW
 XX Homo sapiens.
 OS
 FH Key Location/Qualifiers
 FT Peptide 1..18
 FT /label= presignal
 FT Peptide 19..240
 FT /label= claimed
 FT /note= "as part of fusion polypeptide"
 FT Peptide 19..24
 FT /label= propeptide
 FT Peptide 120..135
 FT /label= claimed
 FT /note= "as part of fusion polypeptide"
 FT
 XX US5408038-A.
 XX
 XX 18-APR-1995.
 XX
 XX 08-OCT-1992; 92US-00959946.
 XX
 XX 09-OCT-1991; 91US-00774633.
 PR
 XX 18-JUN-1992; 92US-00901706.
 XX
 XX (SCRI) SCRIPPS RES INST.
 XX
 XX Witztum JL, Koduri KR, Young SG, Smith RS, Curtiss LK;
 PI
 XX WPI; 1993-134378/16.
 DR
 XX N-PSDB; AAQ89634.
 XX
 XX Polypeptide mimic of native apo B-100 and native apo A-I - useful in
 PT assays for LDL and HDL in plasma samples.
 FT
 XX Claim 10; Fig 2; 41pp; English.
 PS
 XX AA89634 and AAR72705 depict the AA sequence of human apo A-I and its
 CC corresp. cDNA, including presignal residues and propeptide residues,
 CC according to Seilhamer et al., DNA 3(4):309 (1984). A dispersible apo A-
 CC I/B-100 fusion polypeptide is claimed which contains a first AA sequence
 CC of apo A-I and that includes at least AA sequence positions 120-135 (see
 CC AAR72606) and which reacts with anti-apo AI antibodies such as: AI-4
 CC ATCC HB8744; AI-7 ATCC HB 8745; AI-9 ATCC HB 8741; AI-10 ATCC HB 9200; AI
 CC

CC -11 ATCC HB 9201; AI-12 ATCC HB 9202; AI-13 ATCC HB 9203; AI-14 ATCC HB
CC 9204; AI-18 ATCC HB 9507
XX
SQ Sequence 267 AA;

Query Match 100.0%; Score 1362; DB 2; Length 267;
Best Local Similarity 100.0%; Pred. No. 3e-99;
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKAAVLTAVLFTGTSQARHFQQDEPPQSPMDRVKDLATVYVDVKDSGRDYVSQFEGS 60
DB 1 MKAAVLTAVLFTGTSQARHFQQDEPPQSPMDRVKDLATVYVDVKDSGRDYVSQFEGS 60

QY 61 ALGKQLNLKLLDNWDSVTSTFSKLRQLGPGVTQEFWDNLEKETEGRLQEMSKDLEEVKAK 120
DB 61 ALGKQLNLKLLDNWDSVTSTFSKLRQLGPGVTQEFWDNLEKETEGRLQEMSKDLEEVKAK 120

QY 121 VQPYLDDFOKKQOEEMELYRQKVEPLRAELQEGAROKLHELQEKLSPLGEMMDRARAHV 180
DB 121 VQPYLDDFOKKQOEEMELYRQKVEPLRAELQEGAROKLHELQEKLSPLGEMMDRARAHV 180

QY 181 DALRTHLAPYSDELQRLAARLEALKENGARLAEYHAKATEHLSTLSEKAKPALEDLRQ 240
DB 181 DALRTHLAPYSDELQRLAARLEALKENGARLAEYHAKATEHLSTLSEKAKPALEDLRQ 240

QY 241 GLLPVLESFKVSPFLSALEEYTKKLTNQ 267
DB 241 GLLPVLESFKVSPFLSALEEYTKKLTNQ 267

RESULT 4
AAR34032
ID AAR34032 standard; protein; 267 AA.

AC AAR34032;

XX 25-MAR-2003 (revised)

DT 13-AUG-1993 (first entry)

XX Sequence of apo AI.

XX Lipoprotein; apoprotein; B-100; A-I; LDL; HDL; assay.

XX Homo sapiens.

XX WO9307165-A1.

XX 15-APR-1993.

XX 09-OCT-1992; 92WO-US008634.

XX 09-OCT-1991; 91US-00774633.

PR 18-JUN-1992; 92US-00901706.

PR 08-OCT-1992; 92US-00959946.

XX (SCRI) SCRIPPS RES INST.

XX Smith RS, Curtiss LK, Koduri KR, Witztum JL, Young SG;

XX WPI; 1993-134378/16.

DR N-PSDB; AAQ40030.

XX Polypeptide mimic of native apo B-100 and native apo A-I - useful in

PT assays for LDL and HDL in plasma samples.

PT Claim 19; Page 105-106; 137pp; English.

XX The inventors claim a portion of the polypeptide contg. apo B-100 that
CC immunoreacts with antibodies secreted by the hybridoma MB47 having ATCC
CC Accession No. 8746. Polypeptides specifically claimed include residues
CC 217-297, 216-310, 216-331, 216-352, 216-377, 1-377, 205-297, 173-297, 140
CC -297. DNA sequences encoding the polypeptides are also claimed. Also
CC claimed are a fusion polypeptide that contains: (a) a first amino

CC acid residue sequence up to 250 residues in length that includes residues
CC 120-135 of apo A-I, (b) a second amino acid residue sequence up to 375
CC residues in length that includes residues 217-297 of apo B-100 and DNA
CC encoding it. (Updated on 25-MAR-2003 to correct PN field.) (Updated on 25
CC -MAR-2003 to correct PR field.)
XX

SQ Sequence 267 AA;

Query Match 100.0%; Score 1362; DB 2; Length 267;
Best Local Similarity 100.0%; Pred. No. 3e-99;
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKAAVLTAVLFTGTSQARHFQQDEPPQSPMDRVKDLATVYVDVKDSGRDYVSQFEGS 60
DB 1 MKAAVLTAVLFTGTSQARHFQQDEPPQSPMDRVKDLATVYVDVKDSGRDYVSQFEGS 60

QY 61 ALGKQLNLKLLDNWDSVTSTFSKLRQLGPGVTQEFWDNLEKETEGRLQEMSKDLEEVKAK 120
DB 61 ALGKQLNLKLLDNWDSVTSTFSKLRQLGPGVTQEFWDNLEKETEGRLQEMSKDLEEVKAK 120

QY 121 VQPYLDDFOKKQOEEMELYRQKVEPLRAELQEGAROKLHELQEKLSPLGEMMDRARAHV 180
DB 121 VQPYLDDFOKKQOEEMELYRQKVEPLRAELQEGAROKLHELQEKLSPLGEMMDRARAHV 180

QY 181 DALRTHLAPYSDELQRLAARLEALKENGARLAEYHAKATEHLSTLSEKAKPALEDLRQ 240
DB 181 DALRTHLAPYSDELQRLAARLEALKENGARLAEYHAKATEHLSTLSEKAKPALEDLRQ 240

QY 241 GLLPVLESFKVSPFLSALEEYTKKLTNQ 267
DB 241 GLLPVLESFKVSPFLSALEEYTKKLTNQ 267

RESULT 5

AAY18675

ID AAY18675 standard; protein; 267 AA.

XX AAY18675;

XX 09-JUL-1999 (first entry)

XX Human apolipoprotein AI protein sequence.

XX Gene therapy; apolipoprotein A-I; agonist; dyslipidemic disorder; ApoA-I;
KW cardiovascular disease; atherosclerosis; restenosis; LCAT;
KW hyperlipidemia; septic shock; lecithin:cholesterol acyltransferase.

XX Homo sapiens.

XX WO9916409-A2.

XX 08-APR-1999.

XX 28-SEP-1998; 98WO-US020329.

XX 29-SEP-1997; 97US-00940136.

XX (DASS/) DASSEUX J.

PA (SEKU/) SEKUL R.

PA (BUTT/) BUTTNER K.

PA (CORN/) CORNUT I.

PA (METZ/) METZ G.

PA (DUFO/) DUFOURCQ J.

XX Dasseux J, Sekul R, Buttner K, Cornut I, Metz G, Dufourcq J;

XX WPI; 1999-254921/21.

DR N-PSDB; AAX55971.

XX Nucleic acid encoding apoproteinA-I agonist peptides.

XX Example; Fig 1; 232pp; English.

CC The present invention describes a nucleic acid (A) encoding an
 CC apolipoprotein A-I (apoA-I) agonist (B) that is a peptide, or analog,
 CC which forms an amphipathic alpha-helix in presence of lipids. (A),
 CC optionally as a complex with lipids, and host cells that contain (A), are
 CC useful for gene therapy, or prevention, of diseases associated with
 CC dyslipidemia, specifically hypercholesterolemia, cardiovascular disease,
 CC atherosclerosis, restenosis, HDL (high density lipoprotein) and apoA-I
 CC deficiency, hypertriglyceridemia and metabolic syndrome, also to treat
 CC endotoxemia (septic shock). Host cells containing (A) can also be used to
 CC study the role of apoA-I in lipid metabolism. (B) can be used
 CC diagnostically, e.g. to measure serum HDL (particularly its subpopulation
 CC involved in retrograde cholesterol transport) and for imaging the
 CC circulatory system or HDL accumulations at fatty streaks. The present
 CC sequence represents human apoA-I
 XX
 SQ Sequence 267 AA;

Query Match 100.0%; Score 1362; DB 2; Length 267;
 Best Local Similarity 100.0%; Pred. No. 3e-99;
 Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKAAVTLAVLFTGSGARHFQWODEPPQSPWDRVKDLATVYVDVLKDSGRDYVSQFEGS 60
 DB 1 MKAAVTLAVLFTGSGARHFQWODEPPQSPWDRVKDLATVYVDVLKDSGRDYVSQFEGS 60
 QY 61 ALGKQLNLKLDNWDVSTFSLRQLGPVTOEFWDNLEKETEGRLQMSKDLSEVKAK 120
 DB 61 ALGKQLNLKLDNWDVSTFSLRQLGPVTOEFWDNLEKETEGRLQMSKDLSEVKAK 120
 QY 121 VQPYLDDFKKQWQEMELRYQKVEPLRAELQEGARQLHELOKLSPLGEEMRDRARAHV 180
 DB 121 VQPYLDDFKKQWQEMELRYQKVEPLRAELQEGARQLHELOKLSPLGEEMRDRARAHV 180
 QY 181 DALRTHLAPYSDLRQLAARLBALEKNGARLAAYHAKATEHLSTLSEKAPALDRLQ 240
 DB 181 DALRTHLAPYSDLRQLAARLBALEKNGARLAAYHAKATEHLSTLSEKAPALDRLQ 240
 QY 241 GLLPVLESFKVSLSALEEYTKLINTQ 267
 DB 241 GLLPVLESFKVSLSALEEYTKLINTQ 267

RESULT 6
 AAB47620
 ID AAB47620 standard; protein; 267 AA.
 XX
 AC AAB47620;
 DT 21-JAN-2002 (first entry)
 XX
 DE Full length Apo-A1.
 XX
 KW Apolipoprotein; Apo-A1; Apo-A-I fragment T-cell activation inhibitor;
 KW AFT1; monocyte; IL-1; interleukin 1; TNF; tumour necrosis factor;
 KW acute pancreatitis; Alzheimer's disease; asthma; cancer; fever;
 KW inflammatory bowel disease; ischemia; multiple sclerosis; osteoporosis;
 KW Parkinson's disease; psoriasis; probe.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 25..194
 FT /label= AFT1
 FT /note= "18 kD N-terminal fragment"
 FT 25..144
 FT /label= AFT1
 FT /note= "13 kD N-terminal fragment"
 FT 44..65
 FT /label= Helical lipid binding domain
 FT 44..65
 FT /label= Amphipathic helix
 FT 66..120
 FT /label= Phylogenetically conserved domain

FT Domain 66..98
 FT /label= Amphipathic helix
 FT 74..111
 FT /note= "Involved in lipoprotein-mediated cholesterol
 FT efflux from monocytes"
 FT 90..111
 FT /note= "Involved in lectin-cholesterol acyltransferase
 FT activity"
 FT 99..143
 FT /label= Hinged domain
 FT 99..120
 FT /label= Major antigenic epitope domain
 FT 99..120
 FT /label= Amphipathic helix
 FT 121..142
 FT /label= Amphipathic helix
 FT 143..164
 FT /label= Amphipathic helix
 FT 149..219
 FT /label= Receptor binding domain
 FT 156..267
 FT /label= AFT1
 FT /note= "13 kD C-terminal fragment"
 FT 165..208
 FT /label= Amphipathic helix
 FT 209..219
 FT /label= Amphipathic helix
 FT 220..241
 FT /label= Helical lipid binding domain
 FT 220..241
 FT /label= Amphipathic helix
 XX
 PN WO200168852-A2.
 XX
 PD 20-SEP-2001.
 XX
 XX 13-MAR-2001; 2001WO-US007826.
 XX
 XX 13-MAR-2000; 2000US-0189008P.
 XX
 XX (AMGE-) AMGEN INC.
 XX
 XX Edwards CK, Burger D, Dayer J, Kohno T;
 DR WPI; 2001-596908/67.
 DR N-PSDB; AAH43623.
 XX
 PT Apo-A-I fragment T-cell activation inhibitor (AFT1) polynucleotides,
 PT useful for treating, diagnosing, ameliorating diseases associated with IL
 PT -1 and/or TNF activity, e.g. acute pancreatitis, Alzheimer's disease and
 PT asthma.
 XX
 PS Claim 1; Fig 1A; 132pp; English.
 XX
 XX This sequence shows full length apolipoprotein (Apo-A1). Fragments of Apo
 CC -A1 may be used as Apo-A-I fragment T-cell activation inhibitors (AFT1).
 CC These fragments are selected from an 18 kD N-terminal fragment (amino
 CC acids 25-194), a 13 kD N-terminal fragment (amino acids 25-144) and a 13
 CC kD C-terminal fragment (amino acids 156-267). The AFT1 polypeptides and
 CC polynucleotides are useful for regulating T-cell mediated activation of
 CC monocytes and for treating, diagnosing, ameliorating diseases associated
 CC with IL-1 and/or TNF activity. The diseases are acute pancreatitis,
 CC Alzheimer's disease, asthma, cancer, fever, inflammatory bowel disease,
 CC ischemia, multiple sclerosis, osteoporosis, Parkinson's disease or
 CC psoriasis. Numerous examples of other diseases are given in the
 CC specification. The AFT1 nucleic acids are useful as hybridization probes
 CC in diagnostic assays to test for the presence of an AFT1 or Apo-A-I DNA
 CC in mammalian tissue or bodily fluid samples
 XX
 SQ Sequence 267 AA;

Query Match 100.0%; Score 1362; DB 4; Length 267;
 Best Local Similarity 100.0%; Pred. No. 3e-99;

Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKAAVLTAVLFTGQARHFQQDEPPQSPWDRVKDLATVYVDVLKDSGRDYVSQFEGS 60
DB 1 MKAAVLTAVLFTGQARHFQQDEPPQSPWDRVKDLATVYVDVLKDSGRDYVSQFEGS 60

QY 61 ALGKQLNLKLLDNWDSVTSTFSKLRQLGPGVTQEFWDNLEKETEGRLQEMSKDLEEVKAK 120
DB 61 ALGKQLNLKLLDNWDSVTSTFSKLRQLGPGVTQEFWDNLEKETEGRLQEMSKDLEEVKAK 120

QY 121 VQPYLDDFQKKWQOEEMELYRQKVEPLRAELQEGARQKLHLOEKLSPGSEMDRARAHV 180
DB 121 VQPYLDDFQKKWQOEEMELYRQKVEPLRAELQEGARQKLHLOEKLSPGSEMDRARAHV 180

QY 181 DALRTHLAPYSDELRLQRLAARLEALKENGARLAHYHAKATEHLSTLSEKAKPALEDLRQ 240
DB 181 DALRTHLAPYSDELRLQRLAARLEALKENGARLAHYHAKATEHLSTLSEKAKPALEDLRQ 240

QY 241 GLLPVLESFKVSFLSALEYTKKLNQ 267
DB 241 GLLPVLESFKVSFLSALEYTKKLNQ 267

RESULT 7
AAO15892
ID AAO15892 standard; protein; 267 AA.
AC AAO15892;
XT 30-JAN-2003 (first entry)
DE Human apolipoprotein A-1 (APOA1).
KW Human; gene therapy; single nucleotide polymorphism; SNP;
KW cytochrome C oxidase subunit VIb; COX6B; high serum cholesterol; GPI-1;
KW N-acetylglucosaminyl transferase component; cardiovascular disease; HDL;
KW glycosylphosphatidylinositol-1; low serum high density lipoprotein.
OS Homo sapiens.
XX
XX
XX WO200272604-A2.
XX
PD 19-SEP-2002.
XX
XX 05-MAR-2002; 2002WO-US006728.
XX
PR 09-MAR-2001; 2001US-00802640.
XX
XX (SEQU-) SEQUENOM INC.
XX
XX Braun A, Bansal A, Kleyn PW;
XX
XX WPI; 2002-750478/81.
DR N-PSDB; ABT13014.
DR
XX
PT Detecting the presence or absence of an allelic variant of a polymorphic
PT region of COX6B and/or GPI-1 gene, useful for detecting a predisposition
PT to high serum cholesterol, low serum HDL and cardiovascular disease.
XX
XX Disclosure; Page 137-138; 1999p; English.
PS
XX
XX The invention comprises methods of detecting the presence or absence of
CC at least one allelic variant of a polymorphic region of a gene associated
CC with cardiovascular disease. The invention specifically relates to
CC detecting the region of a cytochrome C oxidase subunit VIb (COX6B) gene
CC that is associated with high serum cholesterol, or the region of the N-
CC acetylglucosaminyl transferase component glycosylphosphatidylinositol-1
CC (GPI-1) gene that is associated with low serum high density lipoprotein
CC (HDL). The methods of the invention are useful for detecting a
CC predisposition to high serum cholesterol, low serum HDL and
CC cardiovascular disease. The methods are also useful for elucidating
CC pathological pathways, developing diagnostic assays and new drug
CC therapies for such disorders. The present amino acid sequence represents

CC a human protein that is associated with high serum cholesterol, low serum
XX HDL and/or cardiovascular disease
SQ Sequence 267 AA;

Query Match 100.0%; Score 1362; DB 5; Length 267;
Best Local Similarity 100.0%; Pred. No. 3e-99;
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKAAVLTAVLFTGQARHFQQDEPPQSPWDRVKDLATVYVDVLKDSGRDYVSQFEGS 60
DB 1 MKAAVLTAVLFTGQARHFQQDEPPQSPWDRVKDLATVYVDVLKDSGRDYVSQFEGS 60

QY 61 ALGKQLNLKLLDNWDSVTSTFSKLRQLGPGVTQEFWDNLEKETEGRLQEMSKDLEEVKAK 120
DB 61 ALGKQLNLKLLDNWDSVTSTFSKLRQLGPGVTQEFWDNLEKETEGRLQEMSKDLEEVKAK 120

QY 121 VQPYLDDFQKKWQOEEMELYRQKVEPLRAELQEGARQKLHLOEKLSPGSEMDRARAHV 180
DB 121 VQPYLDDFQKKWQOEEMELYRQKVEPLRAELQEGARQKLHLOEKLSPGSEMDRARAHV 180

QY 181 DALRTHLAPYSDELRLQRLAARLEALKENGARLAHYHAKATEHLSTLSEKAKPALEDLRQ 240
DB 181 DALRTHLAPYSDELRLQRLAARLEALKENGARLAHYHAKATEHLSTLSEKAKPALEDLRQ 240

QY 241 GLLPVLESFKVSFLSALEYTKKLNQ 267
DB 241 GLLPVLESFKVSFLSALEYTKKLNQ 267

RESULT 8
ABG97593
ID ABG97593 standard; protein; 267 AA.
XX
XX AC ABG97593;
XX
DT 17-DEC-2002 (first entry)
XX
XX Human apolipoprotein analogue 1 (Apo A1) protein #15.
XX
XX Apolipoprotein analogue; Apo A; arteriosclerosis; endotoxin removal;
KW angina pectoris; myocardial infarction; arterial stenosis; claudication;
KW carotid stenosis; cerebral arterial stenosis; gene therapy; cholesterol;
KW cardiovascular disease.
XX
XX Homo sapiens.
XX
XX WO200238609-A2.
XX
XX 16-MAY-2002.
XX
XX 09-NOV-2001; 2001WO-DK000739.
XX
XX 10-NOV-2000; 2000DK-00001682.
PR 15-JAN-2001; 2001DK-00000057.
PR 26-JAN-2001; 2001US-0264022P.
XX
XX (PROT-) PROTEOPHARMA APS.
PA
XX
XX Graversen J, Moestrup S;
XX WPI; 2002-527481/56.
DR
XX
XX Novel apolipoprotein construct comprising apolipoprotein A linked to
PT carbohydrate, peptide or protein heterologous group, useful for treating
PT plaque/unstable angina pectoris, myocardial infarction, arterial
PT stenoses.
XX
XX Disclosure; Fig 1; 113pp; English.
XX
XX The invention describes an Apolipoprotein (Apo) construct (I) for use as
CC medicament having general formula apo-A-X, where apo-A is an
CC apolipoprotein component such as apolipoprotein AI, AII or AIV, or its

CC analogue or variant, and X is heterologous group e.g., amino acid,
 CC peptide, protein, carbohydrate or a nucleic acid, providing that when (I)
 CC consists of exactly two identical, native apolipoproteins these are
 CC linked serially. (I) is useful for preparing a pharmaceutical composition
 CC which further comprises excipients, adjuvants, additives, such as
 CC phospholipids, cholesterol or triglycerides. (I) is useful or treating
 CC and/or preventing arteriosclerosis, for removing endotoxins, for treating
 CC angina pectoris including plaque or unstable angina pectoris, myocardial
 CC infarction, arterial stenoses such as claudication, carotis stenosis,
 CC cerebral arterial stenosis and other cardiovascular diseases. The nucleic
 CC acid (II) encoding (I) is useful for gene therapy, where the DNA sequence
 CC encoding (I) is used for transfection or infection of at least one cell
 CC population comprising macrophages or liver cells. (I) has a half-life of
 CC at least the half-life of native apoA-I, A-II or A-IV, preferably two
 CC times higher or more preferably 10 times higher than the half-life of the
 CC apoA molecules. (I) also has a higher binding affinity to cholesterol
 CC compared to native apoA-I, A-II or A-IV. (I) causes substantially no
 CC immune response in humans. This is the amino acid sequence of a human
 CC apolipoprotein (Apo) analogue fragment
 XX
 XX Sequence 267 AA;

Query Match 100.0%; Score 1362; DB 5; Length 267;
 Best Local Similarity 100.0%; Pred. No. 3e-99;
 Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MKAAVLTAVLFLTGQARHFQQDEPPQSPWDRVKDLATVYVDVLDKSGRDYVSQFEGS 60
 DB 1 MKAAVLTAVLFLTGQARHFQQDEPPQSPWDRVKDLATVYVDVLDKSGRDYVSQFEGS 60
 QY 61 ALGKQLNLKLDNDVSTSTFSKLRQLGPVTOEFWDNLEKETEGLRQEMSKDLSEVKAK 120
 DB 61 ALGKQLNLKLDNDVSTSTFSKLRQLGPVTOEFWDNLEKETEGLRQEMSKDLSEVKAK 120
 QY 121 VQPYLDDFQKKQWEEMLYRQKVEPLRAELQEGARQKLHELQEKLSPLGSEMRDRARAHV 180
 DB 121 VQPYLDDFQKKQWEEMLYRQKVEPLRAELQEGARQKLHELQEKLSPLGSEMRDRARAHV 180
 QY 181 DALRTHLAPYSDELRLQRLAARLEALKENGARLAAYHAKATEHLSTLSEKAKPALEDLRQ 240
 DB 181 DALRTHLAPYSDELRLQRLAARLEALKENGARLAAYHAKATEHLSTLSEKAKPALEDLRQ 240
 QY 241 GLLPVLESFKVSFLSALSEETKCLNTQ 267
 DB 241 GLLPVLESFKVSFLSALSEETKCLNTQ 267

RESULT 9
 ABR44031
 ID ABR44031 standard; protein; 267 AA.

XX ABR44031;

AC 04-AUG-2003 (first entry)

XX Human apolipoprotein A-I (ApoA-I) protein.

DE Apolipoprotein A-I; ApoA-I; cardiant; gene therapy; human;
 KW cardiovascular disorder.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 178

FT /label= L178P

FT /note= "wild-type Leu can be replaced with Pro"

XX WO2003035691-A1.

XX 01-MAY-2003.

XX 25-OCT-2002; 2002WO-CA001615.

XX

PR 26-OCT-2001; 2001US-0335075P.
 XX (XENO-) XENON GENETICS INC.
 PA (UYBR-) UNIV BRITISH COLUMBIA.
 XX Brownlie AJ, Dube M, Samuels M, Hayden MR;
 XX WPI: 2003-449258/42.
 DR N-PSDB; ACC48172.
 XX New Apolipoprotein A-1 mutation, useful for preparing a composition for
 PT treating cardiovascular disorder.
 PT Claim 19; Page 33; 52pp; English.
 PS The invention relates to a mutation of the apolipoprotein A-I (ApoA-I)
 CC comprising a leucine to proline variant in position 178. The mutant ApoA-
 CC I polynucleotide is useful for preparing a composition for treating
 CC cardiovascular disorder. The present sequence represents a human ApoA-I
 CC wild-type protein
 XX
 XX Sequence 267 AA;

Query Match 100.0%; Score 1362; DB 6; Length 267;
 Best Local Similarity 100.0%; Pred. No. 3e-99;
 Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MKAAVLTAVLFLTGQARHFQQDEPPQSPWDRVKDLATVYVDVLDKSGRDYVSQFEGS 60
 DB 1 MKAAVLTAVLFLTGQARHFQQDEPPQSPWDRVKDLATVYVDVLDKSGRDYVSQFEGS 60
 QY 61 ALGKQLNLKLDNDVSTSTFSKLRQLGPVTOEFWDNLEKETEGLRQEMSKDLSEVKAK 120
 DB 61 ALGKQLNLKLDNDVSTSTFSKLRQLGPVTOEFWDNLEKETEGLRQEMSKDLSEVKAK 120
 QY 121 VQPYLDDFQKKQWEEMLYRQKVEPLRAELQEGARQKLHELQEKLSPLGSEMRDRARAHV 180
 DB 121 VQPYLDDFQKKQWEEMLYRQKVEPLRAELQEGARQKLHELQEKLSPLGSEMRDRARAHV 180
 QY 181 DALRTHLAPYSDELRLQRLAARLEALKENGARLAAYHAKATEHLSTLSEKAKPALEDLRQ 240
 DB 181 DALRTHLAPYSDELRLQRLAARLEALKENGARLAAYHAKATEHLSTLSEKAKPALEDLRQ 240
 QY 241 GLLPVLESFKVSFLSALSEETKCLNTQ 267
 DB 241 GLLPVLESFKVSFLSALSEETKCLNTQ 267

RESULT 10

ABP57065

ID ABP57065 standard; protein; 267 AA.

XX ABP57065;

XX 14-APR-2003 (first entry)

XX Human apolipoprotein (APOA1) protein SEQ ID NO:10.

XX Atherosclerosis; pro-atherogenic inhibitor; antiatherosclerotic;

KW macrophage-specific expression element; gene therapy; apolipoprotein;

KW APOA1.

XX Homo sapiens.

XX WO2003002061-A2.

XX 09-JAN-2003.

XX 26-JUN-2002; 2002WO-US020350.

XX 26-JUN-2001; 2001US-00893366.

XX (UYSA-) UNIV SAN DIEGO STATE FOUND.

Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 MKAAVLTAVLFLTGSQARHFQODEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGS 60
Db	1 MKAAVLTAVLFLTGSQARHFQODEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGS 60
Qy	61 ALGKQLNLKLDNDWSDVSTFSKLRQLGQPVQTEFWDNLKETEGLRQEMSKDLEEVKAK 120
Db	61 ALGKQLNLKLDNDWSDVSTFSKLRQLGQPVQTEFWDNLKETEGLRQEMSKDLEEVKAK 120
Qy	121 VQPYLDDFKKKWQEEEMELYRQKVEPLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHV 180
Db	121 VQPYLDDFKKKWQEEEMELYRQKVEPLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHV 180
Qy	181 DALRTHLAPYSDELRLAARLEALKENGARLAEVHAKATEHLSTLSEKAKPALEDLRQ 240
Db	181 DALRTHLAPYSDELRLAARLEALKENGARLAEVHAKATEHLSTLSEKAKPALEDLRQ 240
Qy	241 GLLPVLESFKVSLSALEBYTKKLNQ 267
Db	241 GLLPVLESFKVSLSALEBYTKKLNQ 267
RESULT 12	
ADJ83084	
ID	ADJ83084 standard; protein; 267 AA.
XX	
AC	ADJ83084;
XX	
DT	06-MAY-2004 (first entry)
XX	
DE	Human apolipoprotein A-I precursor protein - SEQ ID 75.
XX	
KW	NOVX; cytostatic; antiarteriosclerotic; antidiabetic; antiasthmatic; antiallergic; antiinflammatory; respiratory; antiarthritic;
KW	dermatological; antibacterial; cerebroprotective; vasotropic; cardiant;
KW	haemostatic; hypotensive; hepatotropic; neuroprotective; anorectic;
KW	neotropic; antiulcer; muscular; immunosuppressive; gynaecological;
KW	antipsoriatic; endocrine; ophthalmological; osteopathic;
KW	antiparkinsonian; anticonvulsant; tranquiliser; analgesic; nephrotropic;
KW	antiinfertility; antilipaemic; cardiomyopathy; atherosclerosis; diabetes;
KW	cell signal processing; metabolic pathway; asthma; allergy; emphysema;
KW	autoimmune; graft-versus-host; arthritis; cancer; stroke; haemophilia;
KW	obesity; Alzheimer's; pain; chromosome mapping; tissue typing; human;
KW	apolipoprotein A-I precursor.
XX	
OS	Homo sapiens.
XX	
PN	US2003170630-A1.
XX	
PD	11-SEP-2003.
XX	
PF	21-DEC-2001; 2001US-00032189.
XX	
PR	21-DEC-2000; 2000US-0257495P.
PR	22-DEC-2000; 2000US-0258171P.
PR	20-FEB-2001; 2001US-0269940P.
PR	08-MAR-2001; 2001US-0274192P.
PR	22-MAR-2001; 2001US-0277826P.
PR	29-MAR-2001; 2001US-0279840P.
PR	11-APR-2001; 2001US-0282981P.
PR	13-APR-2001; 2001US-0283656P.
PR	31-JUL-2001; 2001US-0309247P.
PR	10-AUG-2001; 2001US-0311754P.
PR	17-AUG-2001; 2001US-0313331P.
XX	
PA	(ALSO/) ALSOBROOK J P.
PA	(TCHE/) TCHERNEV V T.
PA	(LIUX/) LIU X.
PA	(SPYR/) SPYTER K A.
PA	(ZERR/) ZERRHUSEN B D.
PA	(PATT/) PATTURAJAN M.
PA	(LEPL/) LEPLEY D M.

PA	(BURG/) BURGESS C E.
PA	(SHIM/) SHIMKETS R A.
PA	(GROS/) GROSSE W M.
PA	(SZEK/) SZEKERES E S.
PA	(VERN/) VERNET C A M.
PA	(LIIL/) LI L.
PA	(CASM/) CASMAN S J.
PA	(BOLD/) BOLDOG F L.
PA	(GORM/) GORMAN L.
PA	(GANG/) GANGOLLI E A.
PA	(FERN/) FERNANDES E R.
PA	(RIEG/) RIEGER D K.
PA	(EDIN/) EDINGER S R.
PA	(GUNT/) GUNTHER E.
PA	(MILL/) MILLET I.
PA	(SCIO/) SCIORE P.
PA	(ELLE/) ELLERMAN K.
PA	(MACD/) MACDOUGALL J R.
PA	(SMIT/) SMITHSON G.
XX	
PI	Alsobrook JP, Tchernev VT, Liu X, Spytek KA, Zerhusen BD;
PI	Patturajan M, Lepley DM, Burgess CE, Shimkets RA, Grose WM;
PI	Szekeres ES, Vernet CAM, Li L, Casman SJ, Boldog FL, Gorman L;
PI	Gangolli EA, Fernandes ER, Rieger DK, Edinger SR, Gunther E;
PI	Millet I, Sciore P, Ellerman K, Macdougall JR, Smithson G;
XX	
DR	WPI; 2003-898249/82.
XX	
PT	New NOVX polypeptides and nucleic acid molecules, useful for diagnosing,
PT	preventing or treating NOVX-associated polypeptide disorder, e.g.
PT	cardiomyopathy, atherosclerosis, diabetes, cancer, Parkinson's disease or
PT	asthma.
XX	
PS	Disclosure; SEQ ID NO 75; 263pp; English.
XX	
CC	The invention relates to a novel isolated NOVX polypeptide. The
CC	polypeptide demonstrates cytostatic, antiarteriosclerotic, antidiabetic,
CC	antiasthmatic, antiallergic, antiinflammatory, respiratory,
CC	antiarthritic, dermatological, antibacterial, cerebroprotective,
CC	vasotropic, cardiant, haemostatic, hypotensive, hepatotropic,
CC	neuroprotective, anorectic, neotropic, antiulcer, muscular,
CC	immunosuppressive, gynaecological, antipsoriatic, endocrine,
CC	ophthalmological, osteopathic, antiparkinsonian, anticonvulsant,
CC	tranquilliser, analgesic, nephrotropic, antiinfertility and antilipaemic
CC	activities. The NOVX polypeptide, nucleic acid or antibody of the
CC	invention may be useful for treating or preventing a NOVX-associated
CC	disorder, such as cardiomyopathy, atherosclerosis, diabetes or a disorder
CC	related to cell signal processing and metabolic pathway modulation.
CC	Furthermore, the NOVX polypeptides may be useful for diagnosing, treating
CC	or preventing diseases such as asthma, allergies, emphysema, autoimmune
CC	disease, graft-versus-host disease, arthritis, cancer, stroke,
CC	haemophilia, obesity, Alzheimer's disease and pain. The nucleic acids may
CC	be used as hybridisation probes, in chromosome mapping, tissue typing,
CC	preventive medicine or pharmacogenomics. The current sequence is that of
CC	a protein of the invention which is related to human NOVX protein.
XX	
SQ	Sequence 267 AA;
Query Match 100.0%; Score 1362; DB 7; Length 267;	
Best Local Similarity 100.0%; Pred. No. 3e-99;	
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 MKAAVLTAVLFLTGSQARHFQODEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGS 60
Db	1 MKAAVLTAVLFLTGSQARHFQODEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGS 60
Qy	61 ALGKQLNLKLDNDWSDVSTFSKLRQLGQPVQTEFWDNLKETEGLRQEMSKDLEEVKAK 120
Db	61 ALGKQLNLKLDNDWSDVSTFSKLRQLGQPVQTEFWDNLKETEGLRQEMSKDLEEVKAK 120
Qy	121 VQPYLDDFKKKWQEEEMELYRQKVEPLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHV 180
Db	121 VQPYLDDFKKKWQEEEMELYRQKVEPLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHV 180

Qy	181	DALRTHLAPYSGDELQR	LAARLEALKENG	GARLA	EYHAKA	THEHLSTLS	SEKAP	PALEDLRQ	240
Db	181	DALRTHLAPYSGDELQR	LAARLEALKENG	GARLA	EYHAKA	THEHLSTLS	SEKAP	PALEDLRQ	240
Qy	241	GLLPVLESFKVSF	LSALEEYTKKLNTQ	267					
Db	241	GLLPVLESFKVSF	LSALEEYTKKLNTQ	267					

RESULT 13
ADII19752
ID ADI19752 standard; protein; 267 AA.
XX
XX
AC ADI19752;
XX
XX
DT 22-APR-2004 (first entry)
XX
XX
DE Human APOA1 protein.
XX
XX
KW Pro-atherogenic; atherosclerosis; therapy; human; apolipoprotein; APOA1;
KW macrophage-specific expression element.

OS Homo sapiens.
XX
FN US2004001810-A1.
XX
XX
PD 01-JAN-2004.

AA	
PF	26-JUN-2002; 2002US-00186288.
XX	
PR	26-JUN-2002; 2002US-00186288.

DA
PA (DAVI/) DAVIS R A.
XX
XX
FI Davis RA;
XX
XX WPI; 2004-061641/06.
DR N-PSDE; ADI19751.
DR GENBANK; XM 006435.
DR

PT New nucleic acids encoding an inhibitor of a pro-atherogenic molecule
PT operationally linked to a macrophage-specific expression element, useful
PT for reducing or inhibiting atherosclerosis.

PS Disclosure; SEQ ID NO 10; 45pp; English.

CC The present invention provides a nucleic acid comprising a nucleotide
CC sequence encoding an inhibitor of a pro-atherogenic molecule
CC operationally linked to a macrophage-specific expression element. The
CC invention is useful for reducing and inhibiting atherosclerosis. The
CC present sequence is human apolipoprotein (APOA1) protein.

AA	Sequence 267	AA
SQ	Sequence 267	AA
	Query Match	100.0%
	Score	1362
	DB	8
	Length	267

```

Best local similarity 100.0%; Fied. NO. 38-33;
Matches 287; Conservative 0; Mismatches 0; Indels 0; Gaps 0
QY 1 MKAAVLTILAVLFTGSGARHFQQDEPPQSPQPDWRVKDLATVYDVLKSGRDYVVSQFEGS 60

```

Db 1 MKAAVLTAVLFLTGSQARHFWQDEPPQSPWDRVKDLATVYVDVLKDSGRDYVSQFEGS 60

Qy	61	ALGQNLNKLDDNDVSVTFFSKLRQLGPTQEFWDNLKEKETEGIRQEMSKDLEVKAK	120
	61	ALGQNLNKLDDNDVSVTFFSKLRQLGPTQEFWDNLKEKETEGIRQEMSKDLEVKAK	120
Db	61	ALGQNLNKLDDNDVSVTFFSKLRQLGPTQEFWDNLKEKETEGIRQEMSKDLEVKAK	120
	61	ALGQNLNKLDDNDVSVTFFSKLRQLGPTQEFWDNLKEKETEGIRQEMSKDLEVKAK	120

Qy	121	VQPYLDDFQKKWQEMELTRYQKVESPLRAELQEGARQKHELQEKLSPLGEMNRDRARAHV	180
	121	VQPYLDDFQKKWQEMELTRYQKVESPLRAELQEGARQKHELQEKLSPLGEMNRDRARAHV	180
Db	121	VQPYLDDFQKKWQEMELTRYQKVESPLRAELQEGARQKHELQEKLSPLGEMNRDRARAHV	180
	121	VQPYLDDFQKKWQEMELTRYQKVESPLRAELQEGARQKHELQEKLSPLGEMNRDRARAHV	180

QY 181 DALRTHLAPYSDELRQLAARLEALKENGCGARLAEYHAKATEHLSLTSSEKAPALDRLQ 240

Db	181	DAURTHLAPYSDLRQRLAARLEALKENGARLAFYHAKATEHLSTLSEKAKPALEDLRQ	240
Qy	241	GLLPVLESFKVSFLSALEEYTKKLTNQ	267
Db	241	GLLPVLESFKVSFLSALEEYTKKLTNQ	267

RESULT 14	
AAO30162	
ID	AAO30162 standard; protein; 275 AA.
XX	
AC	AAO30162;
XX	
DT	03-SEP-2003 (first entry)
XX	
DE	Human APOA1 protein.
XX	
KW	Human; aging-associated disease; oxidative stress. AAD; APOA1.

XX
OS
Homo sapiens.

AA PN WO2003045988-A2.

PD 05-JUN-2003.

PF 28-NOV-2002; 2002WO-EP013549.

PR 29-NOV-2001; 2001EP-00204500.

PA (VLA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.

PI Contreras RH, Chen C;
vv

DR WPI; 2003-505183/47.
DB N. DODD. 20150805

XX
PT screening genes involved in aging and/or aging-associated diseases

PT oxidative stress by mutating and cultivating a yeast cell, enr

PT population: labeling the mother cells and isolating the highly

PT
cells.
XX

PS CLAIM 13; Page 40; /8pp; English.
XX

CC and/or aging-associated diseases (AAD) or in oxidative stress.
CC the invention relates to a method for screening genes involved in

CC for mother cells, labelling the mother cells and isolating the CC involves mutagenesis and culturing a yeast cell, ensuring the

sequence is used to illustrate the method of the invention

SQ Sequence 275 AA;

Query Match 100.0%; Score 1362; DB 6; Length 275;

Matches 267; Conservative 0; Mismatches 0; Indels 0;

Qy 1 MKAAVLTLAVLFLTGSQARHFQQDEPPQSPWDRVKDLATVYVDVLKDSGRDIY

Db 9 MKA AVL TLAVL FLTGS QARHFQQDEPPQSPWDRVKDLATVYVDVLKDSGRDI

61 ALGKQLNKLNDWDSVTSTFSKLRQLGPVTQFFWDNLLEKETGLRQMSR
QY

DB 69 ALGKQLNKLNDNWDVSVISTFSKLRQQLGPPVIQEFWDNLKRETEGLKQEMSRD

QY IZI VQPI LDDF QKNQ EEMELI RQKV EPRK AEL VEGAR QALH ELX EKT SFL DGEEM

DD VFIDDDFQANWQESMSIIKQVBFUNAHETJEGARKQNGHBJUKRUSFDGEENT

[illegible][illegible]

THE UNIVERSITY OF CHICAGO

Db 249 GLLPVLESFKVSFLSALEYTKKLNQ 275
RESULT 15
ADJ68447
ID ADJ68447 standard; protein; 267 AA.
XX AC ADJ68447;
XX
XX 06-MAY-2004 (first entry)
XX Human heat mitochondrial protein as a therapeutic target SeqID253.
XX
XX mitochondrial; human; screening assay; diabetes mellitus;
KW Huntington's disease; osteoarthritis;
KW Leber's hereditary optic neuropathy; LHON;
KW mitochondrial encephalopathy lactic acidosis and stroke; MELAS;
KW myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;
KW neuroprotective; nontropic; antidiabetic; anticonvulsant; antiarthritic;
KW osteopathic; ophthalmological; cytostatic.
XX
XX Homo sapiens.
XX OS
XX WO2003087768-A2.
XX PN
XX
XX 23-OCT-2003.
XX PD
XX
XX 04-APR-2003; 2003WO-US010870.
XX PF
XX 12-APR-2002; 2002US-0372843P.
XX PR
XX 17-JUN-2002; 2002US-0389987P.
XX PR
XX 20-SEP-2002; 2002US-0412418P.
XX PA
XX (MITO-) MITOKOR.
XX PA
XX (BUCK-) BUCK INST AGE RES.
XX
XX Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM;
PI Warnock DE;
PI
XX WPI; 2003-845369/78.
XX
XX Identifying a mitochondrial target for drug screening assays and for
PT treating diseases associated with altered mitochondrial function,
PT comprises detecting a modified polypeptide in a sample and correlating
PT with the disease.
XX
XX Claim 1; SEQ ID NO 253; 180pp; English.
XX
XX This invention relates to novel mitochondrial targets that can be used
CC for therapeutic intervention in treating a disease associated with
CC altered mitochondrial function. Specifically, it refers to a method for
CC identifying proteins of the human heart mitochondrial proteome that are
CC useful for drug screening assays, as well as therapeutic targets. The
CC present invention describes a method for identifying such proteins that
CC can be used in the treatment of various diseases associated with altered
CC mitochondrial function including diabetes mellitus, Huntington's disease,
CC osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial
CC encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy
CC ragged red fibre syndrome (MERRF) or cancer. Accordingly, these
CC compositions have neuroprotective, nontropic, antidiabetic,
CC anticonvulsant, antiarthritic, osteopathic, ophthalmological and
CC cytotatic activities. This polypeptide sequence is a human heart
CC mitochondrial protein of the invention.
XX
XX Sequence 267 AA;
SQ

Query Match 99.8%; Score 1359; DB 7; Length 267;
Best Local Similarity 99.8%; Pred. No. 5.3e-99;
Matches 266; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 MKA AVLTLAVLFTGSOARHFQQDEPPQSPWDRVKDLATVYVDVLKDSGRDYVSQFEGS 60
|||||

Db 1 MKA AVLTLAVLFTGSOARHFQQDEPPQSPWDRVKDLATVYVDVLKDSGRDYVSQFEGS 60
QY 61 ALGKQLNLKLLDNWDSVTSTFSKLRQLGPTQEFWDNLEKETEGLRQEMSKDLEEVKAK 120
|||||
Db 61 ALGKQLNLKLLDNWDSVTSTFSKLRQLGPTQEFWDNLEKETEGLRQEMSKDLEEVKAK 120
|||||
QY 121 VQPYLDDDFQKKWQEEEMELYRQKVEPLRAELOEGARQKLHELOEKLSPLGSEEMRDRARAHV 180
|||||
Db 121 VQPYLDDDFQKKWQEEEMELYRQKVEPLRAELOEGARQKLHELOEKLSPLGSEEMRDRARAHV 180
|||||
QY 181 DALRTHLAPYSDELRFQRLAARLEALKENGARLAEYHAKATEHLSTLSEKAKPALEDLRQ 240
|||||
Db 181 DALRTHLAPYSDELRFQRLAARLEALKENGARLAEYHAKATEHLSTLSEKAKPALEDLRQ 240
|||||
QY 241 GLLPVLESFKVSFLSALEYTKKLNQ 267
|||||
Db 241 GLLPVLESFKVSFLSALEYTKKLNQ 267
|||||
Search completed: December 21, 2004, 07:21:35
Job time : 163.744 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2004, 06:52:14 ; Search time 132.033 Seconds
(without alignments)
723.622 Million cell updates/sec

Title: US-09-803-918A-2

Perfect score: 1362

Sequence: 1 MKAAVLTAVLFLTGSQARH.....SPKVSFLSALEYTKKLNTQ 267

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1589859 seqs, 357834939 residues

Total number of hits satisfying chosen parameters: 1589859

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1362	100.0	267	9	US-09-803-918A-2
2	1362	100.0	267	9	US-09-803-918A-2
3	1362	100.0	267	10	US-09-802-640-30
4	1362	100.0	267	14	US-10-032-189-75
5	1362	100.0	267	14	US-10-283-599-260
6	1362	100.0	267	14	US-10-403-902A-30
7	1362	100.0	267	15	US-10-186-288-10
8	1362	100.0	267	15	US-10-038-854-219
9	1362	100.0	267	15	US-10-038-854-226
10	1359	99.8	267	14	US-10-032-189-74
11	1359	99.8	267	16	US-10-408-765A-253
12	1346	98.8	267	10	US-09-919-039-27
13	1306	95.9	267	9	US-09-987-107-16

14	1306	95.9	267	14	US-10-032-189-77	Sequence 77, Appl
15	1306	95.9	267	15	US-10-038-854-223	Sequence 223, App
16	1306	95.9	267	15	US-10-038-854-230	Sequence 230, App
17	1299	95.4	267	14	US-10-032-189-78	Sequence 78, Appl
18	1281	94.1	249	14	US-10-032-189-76	Sequence 76, Appl
19	1281	94.1	249	15	US-10-038-854-220	Sequence 220, App
20	1281	94.1	249	15	US-10-038-854-227	Sequence 227, App
21	1276	93.7	250	15	US-10-465-789A-2	Sequence 22, Appl
22	1246.5	91.5	306	9	US-09-987-107-9	Sequence 9, Appl
23	1246	91.5	304	9	US-09-987-107-7	Sequence 7, Appl
24	1246	91.5	304	9	US-09-987-107-8	Sequence 8, Appl
25	1246	91.5	323	9	US-09-987-107-58	Sequence 58, Appl
26	1246	91.5	323	9	US-09-987-107-60	Sequence 60, Appl
27	1243.5	91.3	304	9	US-09-987-107-6	Sequence 6, Appl
28	1243.5	91.3	306	9	US-09-987-107-10	Sequence 10, Appl
29	1243.5	91.3	306	9	US-09-987-107-11	Sequence 11, Appl
30	1243.5	91.3	323	9	US-09-987-107-56	Sequence 56, Appl
31	1242.5	91.2	337	9	US-09-987-107-46	Sequence 46, Appl
32	1242	91.2	301	9	US-09-987-107-3	Sequence 3, Appl
33	1242	91.2	301	9	US-09-987-107-5	Sequence 5, Appl
34	1242	91.2	316	9	US-09-987-107-48	Sequence 48, Appl
35	1242	91.2	316	9	US-09-987-107-54	Sequence 54, Appl
36	1242	91.2	329	9	US-09-987-107-14	Sequence 14, Appl
37	1242	91.2	344	9	US-09-987-107-68	Sequence 68, Appl
38	1241	91.1	243	9	US-09-987-107-1	Sequence 1, Appl
39	1241	91.1	243	15	US-10-038-854-221	Sequence 221, App
40	1241	91.1	243	15	US-10-038-854-228	Sequence 228, App
41	1241	91.1	244	9	US-09-987-107-2	Sequence 2, Appl
42	1241	91.1	261	9	US-09-987-107-52	Sequence 52, Appl
43	1241	91.1	336	9	US-09-987-107-44	Sequence 44, Appl
44	1238	90.9	324	9	US-09-987-107-62	Sequence 62, Appl
45	1235	90.7	324	9	US-09-987-107-64	Sequence 64, Appl

ALIGNMENTS

RESULT 1

US-09-803-918A-2
; Sequence 2, Application US/09803918A
; Patent No. US20020064820A1
; GENERAL INFORMATION:
; APPLICANT: Dayer, Jean-Michel
; APPLICANT: Burger, Danielle
; APPLICANT: Kohno, Tadahiko
; APPLICANT: Edwards III, Carl K.
; TITLE OF INVENTION: APO-A-1 REGULATION OF T-CELL SIGNALING
; FILE REFERENCE: 06843.0035-00000
; CURRENT APPLICATION NUMBER: US/09/803,918A
; CURRENT FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: 60/189,008
; PRIOR FILING DATE: 2000-03-13
; PRIOR APPLICATION NUMBER: 60/193,551
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 267
; TYPE: PRY
; ORGANISM: Homo sapiens
; US-09-803-918A-2

Query Match	100.0%	Score 1362;	DB 9;	Length 267;
Best Local Similarity	100.0%	Pred. No. 1.2e-92;		
Matches 267;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MKAAVLTAVLFLTGSQARHFQODEPPQSPMDRVKDLATVYVDVLKDSGRDYVSQFEGS	60	
Db	1	MKAAVLTAVLFLTGSQARHFQODEPPQSPMDRVKDLATVYVDVLKDSGRDYVSQFEGS	60	
QY	61	ALGKQLNLKLLDNWDSVTSTFSKLRHQLGPVTOBFWNDLEKETEGLRQEMSKDLSEVKAK	120	
Db	61	ALGKQLNLKLLDNWDSVTSTFSKLRHQLGPVTOBFWNDLEKETEGLRQEMSKDLSEVKAK	120	

Qy	121	VQPYLDDFOKKQWQEMELYRQKVPEPLRAELQEGARQKLHELQEKLSPLGEMDRARAHV	180
Db	121	VQPYLDDFOKKQWQEMELYRQKVPEPLRAELQEGARQKLHELQEKLSPLGEMDRARAHV	180
Qy	181	DALRTHLAPYSDELQRLLAARLEALKENGGRARLAEYHAKATEHJLSTLSEKAKPALBDLRQ	240
Db	181	DALRTHLAPYSDELQRLLAARLEALKENGGRARLAEYHAKATEHJLSTLSEKAKPALBDLRQ	240
Qy	241	GLLPVLESFKVSLSALEYTKKLNQ	267
Db	241	GLLPVLESFKVSLSALEYTKKLNQ	267

RESULT 2

US-09-987-107-15

```

; Sequence 15, Application US/09987107
; Patent No. US20020156007A1
; GENERAL INFORMATION:
; APPLICANT: GRAVERSEN, Jonas
; APPLICANT: MOESTRUP, Soren
; TITLE OF INVENTION: APOLIPOPROTEINS ANALOGUES
; FILE REFERENCE: GRAVERSENIA
; CURRENT APPLICATION NUMBER: US/09/987,107
; CURRENT FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/264,022
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: DK PA2001 00057
; PRIOR FILING DATE: 2001-01-15
; PRIOR APPLICATION NUMBER: DK PA2000 01692
; PRIOR FILING DATE: 2000-11-10
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 15
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-987-107-15

```

Query Match 100.0%; Score 1362; DB 9; Length 267;
Best Local Similarity 100.0%; Pred. No. 1.2e-92;
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

[illegible]

RESULT 3

US-09-802-640-30
; Sequence 30, Application US/09802640
; Publication No. US20030036057A1
; GENERAL INFORMATION:
; APPLICANT: Braun, Andreas
; APPLICANT: Bonsal Aruna
; APPLICANT: Kleya Patrick
; TITLE OF INVENTION: GENES AND POLYMORPHISMS ASSOCIATED WITH

;
; PRIOR APPLICATION NUMBER: 60/257,495
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/258,171
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 60/269,940
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/274,192
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/277,826
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: 60/279,840
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/282,981
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 60/283,656
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/309,247
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: 60/311,754
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: 60/313,331
; PRIOR FILING DATE: 2001-08-17
; NUMBER OF SEQ ID NOS: 260
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 75
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-032-189-75

Query Match 100.0%; Score 1362; DB 14; Length 267;
Best Local Similarity 100.0%; Pred. No. 1.2e-92;
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKAAVLTAVLFTGSGARHFQWDEPPQSPMDRVKDLATVYVDVLKDSGRDVSQPEGS 60
DB 1 MKAAVLTAVLFTGSGARHFQWDEPPQSPMDRVKDLATVYVDVLKDSGRDVSQPEGS 60

QY 61 ALGKQLNLKLLDNWDSVTSTFSKLRQGLGPTQEFWDNLEKETEGLRQEMSKDLEEVKAK 120
DB 61 ALGKQLNLKLLDNWDSVTSTFSKLRQGLGPTQEFWDNLEKETEGLRQEMSKDLEEVKAK 120

QY 121 VQPYLDDFQKKWQOEEMELYRQKVEPLRAELQEGARQKLHLOKLSPLGSEMDRARAHV 180
DB 121 VQPYLDDFQKKWQOEEMELYRQKVEPLRAELQEGARQKLHLOKLSPLGSEMDRARAHV 180

QY 181 DALRTHLAPYSDELQRLAARLEALKENGARLAAYHAKATEHLSTLSEKAKPALEDLRQ 240
DB 181 DALRTHLAPYSDELQRLAARLEALKENGARLAAYHAKATEHLSTLSEKAKPALEDLRQ 240

QY 241 GLLPVLESFKVSFLSALEEYTKKLTNQ 267
DB 241 GLLPVLESFKVSFLSALEEYTKKLTNQ 267

RESULT 5
US-10-283-599-260
; Sequence 260, Application US/10283599
; Publication No. US20030208059A1
; GENERAL INFORMATION:
; APPLICANT: Daseux, Jean-Louis
; APPLICANT: Sekul, Renate
; APPLICANT: Buttner, Klaus
; APPLICANT: Cornut, Isabelle
; APPLICANT: Metz, Gunther
; APPLICANT: Dufourcq, Jean
; TITLE OF INVENTION: GENE THERAPY APPROACHES TO
; TITLE OF INVENTION: SUPPLY APOIPOPOTEPIN A-I AGONISTS AND THEIR
; TITLE OF INVENTION: USE TO TREAT DYSLIPIDEMIC DISORDERS.
; NUMBER OF SEQUENCES: 274
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas

;
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/283,599
; FILING DATE: 29-OCT-2002
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/940,136
; FILING DATE: 29-SEP-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 009196-0007-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-493-4935
; TELEFAX: 650-493-5556
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 260:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 267 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
US-10-283-599-260

Query Match 100.0%; Score 1362; DB 14; Length 267;
Best Local Similarity 100.0%; Pred. No. 1.2e-92;
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKAAVLTAVLFTGSGARHFQWDEPPQSPMDRVKDLATVYVDVLKDSGRDVSQPEGS 60
DB 1 MKAAVLTAVLFTGSGARHFQWDEPPQSPMDRVKDLATVYVDVLKDSGRDVSQPEGS 60

QY 61 ALGKQLNLKLLDNWDSVTSTFSKLRQGLGPTQEFWDNLEKETEGLRQEMSKDLEEVKAK 120
DB 61 ALGKQLNLKLLDNWDSVTSTFSKLRQGLGPTQEFWDNLEKETEGLRQEMSKDLEEVKAK 120

QY 121 VQPYLDDFQKKWQOEEMELYRQKVEPLRAELQEGARQKLHLOKLSPLGSEMDRARAHV 180
DB 121 VQPYLDDFQKKWQOEEMELYRQKVEPLRAELQEGARQKLHLOKLSPLGSEMDRARAHV 180

QY 181 DALRTHLAPYSDELQRLAARLEALKENGARLAAYHAKATEHLSTLSEKAKPALEDLRQ 240
DB 181 DALRTHLAPYSDELQRLAARLEALKENGARLAAYHAKATEHLSTLSEKAKPALEDLRQ 240

QY 241 GLLPVLESFKVSFLSALEEYTKKLTNQ 267
DB 241 GLLPVLESFKVSFLSALEEYTKKLTNQ 267

RESULT 6
US-10-403-902A-30
; Sequence 30, Application US/10403902A
; Publication No. US2003022418A1
; GENERAL INFORMATION:
; APPLICANT: Bansal, Aruna
; APPLICANT: Klevy, Patrick
; TITLE OF INVENTION: GENES AND POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISEASE AND THEIR USE
; FILE REFERENCE: 24736-2048B
; CURRENT APPLICATION NUMBER: US/10/403,902A
; CURRENT FILING DATE: 2003-07-21
; PRIOR APPLICATION NUMBER: 09/802,640

```
; PRIOR FILING DATE: 2001-03-09
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-403-902A-30

Query Match      100.0%; Score 1362; DB 14; Length 267;
Best Local Similarity 100.0%; Pred. No. 1.2e-92;
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKAAVLTAVLFLTGSQARHFHQDEPPQSPMDRVKDLATVYVDVLKDSGRDVVSQFEGS 60
DB 1 MKAAVLTAVLFLTGSQARHFHQDEPPQSPMDRVKDLATVYVDVLKDSGRDVVSQFEGS 60
QY 61 ALGKQLNLKLLDNWDSVTSTFSLRQLGPGVTQGFWDNLEKEKEGRLQEMSKDLEEVKAK 120
DB 61 ALGKQLNLKLLDNWDSVTSTFSLRQLGPGVTQGFWDNLEKEKEGRLQEMSKDLEEVKAK 120
QY 121 VQPYLDDFOKKQOEEMELVYRQKVEPLRAELQEGAROKLHELQBLQKLSPLGGEEMDRARAHV 180
DB 121 VQPYLDDFOKKQOEEMELVYRQKVEPLRAELQEGAROKLHELQBLQKLSPLGGEEMDRARAHV 180
QY 181 DALRTHLAPYSDELRLAARLEALKENGARLAEYHAKATEHLSTLSEKAKPALEDLRQ 240
DB 181 DALRTHLAPYSDELRLAARLEALKENGARLAEYHAKATEHLSTLSEKAKPALEDLRQ 240
QY 241 GLLPVLESFKVSFLSALEYTKKLTNQ 267
DB 241 GLLPVLESFKVSFLSALEYTKKLTNQ 267

RESULT 7
US-10-186-288-10
; Sequence 10, Application US/10186288
; Publication No. US20040001810A1
; GENERAL INFORMATION:
; APPLICANT: Davis, Roger A.
; TITLE OF INVENTION: Compositions and Methods for Treating
; FILE REFERENCE: P-RD 5299
; CURRENT APPLICATION NUMBER: US/10/186,288
; CURRENT FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: US 09/893,366
; PRIOR FILING DATE: 2001-06-26
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-186-288-10

Query Match      100.0%; Score 1362; DB 15; Length 267;
Best Local Similarity 100.0%; Pred. No. 1.2e-92;
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKAAVLTAVLFLTGSQARHFHQDEPPQSPMDRVKDLATVYVDVLKDSGRDVVSQFEGS 60
DB 1 MKAAVLTAVLFLTGSQARHFHQDEPPQSPMDRVKDLATVYVDVLKDSGRDVVSQFEGS 60
QY 61 ALGKQLNLKLLDNWDSVTSTFSLRQLGPGVTQGFWDNLEKEKEGRLQEMSKDLEEVKAK 120
DB 61 ALGKQLNLKLLDNWDSVTSTFSLRQLGPGVTQGFWDNLEKEKEGRLQEMSKDLEEVKAK 120
QY 121 VQPYLDDFOKKQOEEMELVYRQKVEPLRAELQEGAROKLHELQBLQKLSPLGGEEMDRARAHV 180
DB 121 VQPYLDDFOKKQOEEMELVYRQKVEPLRAELQEGAROKLHELQBLQKLSPLGGEEMDRARAHV 180
QY 181 DALRTHLAPYSDELRLAARLEALKENGARLAEYHAKATEHLSTLSEKAKPALEDLRQ 240
DB 181 DALRTHLAPYSDELRLAARLEALKENGARLAEYHAKATEHLSTLSEKAKPALEDLRQ 240
```

```
DB 181 DALRTHLAPYSDELRLAARLEALKENGARLAEYHAKATEHLSTLSEKAKPALEDLRQ 240
QY 241 GLLPVLESFKVSFLSALEYTKKLTNQ 267
DB 241 GLLPVLESFKVSFLSALEYTKKLTNQ 267

RESULT 8
US-10-038-854-219
; Sequence 219, Application US/10038854
; Publication No. US20040022781A1
; GENERAL INFORMATION:
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Li, Li
; APPLICANT: Wolenc, Adam R
; APPLICANT: Vernet, Corine
; APPLICANT: Eisen, Andrew J
; APPLICANT: Liu, Xiaohong
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Shinkets, Richard A
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spaderma, Steven K
; APPLICANT: Gorman, Linda
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Gangolli, Basha A
; APPLICANT: Guo, Xiaojia S
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Rastelli, Luca
; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc
; APPLICANT: Burgess, Catherine E
; APPLICANT: Edinger, Shalomit R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gunther, Erik
; APPLICANT: Smithson, Glennnda
; APPLICANT: Millet, Isabelle
; APPLICANT: MacDougall, John R
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-230
; CURRENT APPLICATION NUMBER: US/10/038,854
; CURRENT FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/258,928
; PRIOR FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: 60/259,415
; PRIOR FILING DATE: 2001-01-02
; PRIOR APPLICATION NUMBER: 60/259,785
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/269,814
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/279,832
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,833
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,863
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/283,889
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,447
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/286,683
; PRIOR FILING DATE: 2001-04-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 219
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-038-854-219

Query Match      100.0%; Score 1362; DB 15; Length 267;
```

[illegible]

```

RESULT 9
US-10-038-854-226
; Sequence 226, Application US/10038854
; Publication No. US20040022781A1
; GENERAL INFORMATION:
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Li, Li
; APPLICANT: Wolenc, Adam R
; APPLICANT: Vernet, Corine
; APPLICANT: Eisen, Andrew J
; APPLICANT: Liu, Xiaohong
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Shmkets, Richard A
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spaderna, Steven K
; APPLICANT: Gorman, Linda
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Gangolli, Esha A
; APPLICANT: Guo, Xiaojia S
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Rastelli, Luca
; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc
; APPLICANT: Burgess, Catherine E
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gunther, Erik
; APPLICANT: Smithson, Glennda
; APPLICANT: Millet, Isabelle
; APPLICANT: MacDougall, John R
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-230
; CURRENT APPLICATION NUMBER: US/10/038,854
; CURRENT FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/258,928
; PRIOR FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: 60/259,415
; PRIOR FILING DATE: 2001-01-02
; PRIOR APPLICATION NUMBER: 60/259,785
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/269,814
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/279,832
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,833
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,863

```

```

; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/283,889
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,447
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/286,683
; PRIOR FILING DATE: 2001-04-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 226
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-038-854-226

Query Match      100.0%; Score 1362; DB 15; Length 267;
Best Local Similarity 100.0%; Pred. No. 1.2e-92;
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKAAVLTAVLFTGSGARHFHQDEPPQSPWDRVKDLATVYDVLDKSDGRDYVSQPEGS 60
Db 1 MKAAVLTAVLFTGSGARHFHQDEPPQSPWDRVKDLATVYDVLDKSDGRDYVSQPEGS 60

Qy 61 ALGKQLNLKLLDNWDSVTSTFSKLREQLGVPVTOEFWDNLEKETEGLEQMSKDLEEVKAK 120
Db 61 ALGKQLNLKLLDNWDSVTSTFSKLREQLGVPVTOEFWDNLEKETEGLEQMSKDLEEVKAK 120

Qy 121 VQPVLDQFKKQGEEMELYKQKEVPLRAELQEGARQKLHLEQKLSPLGEMRDRARAHV 180
Db 121 VQPVLDQFKKQGEEMELYKQKEVPLRAELQEGARQKLHLEQKLSPLGEMRDRARAHV 180

Qy 181 DALRTHLAPYSDELQRFLAARLEALKENGGARLAEYHAKATEHLSTLSEKAKPALEDLRQ 240
Db 181 DALRTHLAPYSDELQRFLAARLEALKENGGARLAEYHAKATEHLSTLSEKAKPALEDLRQ 240

Qy 241 GLLPVLESFKVFSLSALEEYTKKLTQ 267
Db 241 GLLPVLESFKVFSLSALEEYTKKLTQ 267

RESULT 10
US-10-032-189-74
; Sequence 74, Application US/10032189
; Publication No. US20030170630A1
; GENERAL INFORMATION:
; APPLICANT: Alsbrook II, John P
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Liu, Xiaohong
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Patturajan, Meera
; APPLICANT: Grosse, William M
; APPLICANT: Lepley, Denise M
; APPLICANT: Burgees, Catherine E
; APPLICANT: Shimkets, Richard A
; APPLICANT: Grosse, William M
; APPLICANT: Szekeres, Edward S
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Li, Li
; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc L
; APPLICANT: Gorman, Linda
; APPLICANT: Gangolli, Esha A
; APPLICANT: Fernandes, Elma R
; APPLICANT: Rieger, Daniel K
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Gunther, Erik
; APPLICANT: Millet, Isabelle
; APPLICANT: Sciore, Paul
; APPLICANT: Ellerman, Karen
; APPLICANT: MacDougall, John R
; APPLICANT: Smithson, Glennda

```

RESULT 10

US-10-032-189-74

Sequence 74, Application US/10032189
Publication No. US20030170630A1
GENERAL INFORMATION:
APPLICANT: Alsbrook II, John P
APPLICANT: Tchernev, Velizar T
APPLICANT: Liu, Xiaohong
APPLICANT: Spytek, Kimberly A
APPLICANT: Zerhusen, Bryan D
APPLICANT: Patturajan, Meera
APPLICANT: Grosse, William M
APPLICANT: Lepley, Denise M
APPLICANT: Burgess, Catherine E
APPLICANT: Shimkets, Richard A
APPLICANT: Grosse, William M
APPLICANT: Szekeres, Edward S
APPLICANT: Vernet, Corine A.M.
APPLICANT: Li, Li
APPLICANT: Casman, Stacie J
APPLICANT: Boldog, Ferenc L
APPLICANT: Gorman, Linda
APPLICANT: Gangolli, Esha A
APPLICANT: Fernandes, Elma R
APPLICANT: Rieger, Daniel K
APPLICANT: Edinger, Shlomit R
APPLICANT: Gunther, Erik
APPLICANT: Millet, Isabelle
APPLICANT: Sciore, Paul
APPLICANT: Ellerman, Karen
APPLICANT: MacDougall, John R
APPLICANT: Smithson, Glennda

```
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-228
; CURRENT APPLICATION NUMBER: US/10/032,189
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/257,495
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/258,171
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 60/269,940
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/274,192
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/277,826
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: 60/279,840
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/282,981
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 60/283,656
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/309,247
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: 60/311,754
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: 60/313,331
; PRIOR FILING DATE: 2001-08-17
; NUMBER OF SEQ ID NOS: 260
; SOFTWARE: Patent.in Ver. 2.1
; SEQ ID NO 74
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-032-189-74

Query Match          99.8%; Score 1359; DB 14; Length 267;
Best Local Similarity 99.6%; Pred. No. 1.9e-92;
Matches 266; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKAAVLTTLAVLFLTGSQARHFQQDEPPQSPMDRVKDLATVYVDVLKDSGRDYVSQFEGS 60
Db 1 MKAAVLTTLAVLFLTGSQARHFQQDEPPQSPMDRVKDLATVYVDVLKDSGRDYVSQFEGS 60
QY 61 ALGKQLNLKLDNWDVSTFTSKLRQLGPVTQEFWDNLEKETEGLRQEMSKDLBEVKAK 120
Db 61 ALGKQLNLKLDNWDVSTFTSKLRQLGPVTQEFWDNLEKETEGLRQEMSKDLBEVKAK 120
QY 121 VQPYLDDFQKKWQEEEMELYRQKVEPLRAELQEGARQKLHELOEKLSPGLGEEMRDRARAHV 180
Db 121 VQPYLDDFQKKWQEEEMELYRQKVEPLRAELQEGARQKLHELOEKLSPGLGEEMRDRARAHV 180
QY 181 DALRTHLAPYSDELQRLAARLEALKENGARLAELYHAKATEHLSTLSEKAKPALEDLRQ 240
Db 181 DALRTHLAPYSDELQRLAARLEALKENGARLAELYHAKATEHLSTLSEKAKPALEDLRQ 240
QY 241 GLLPVLESFKVSFLSALEYTKKLNQ 267
Db 241 GLLPVLESFKVSFLSALEYTKKLNQ 267
```

```
RESULT 11
US-10-408-765A-253
; Sequence 253, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
```

```
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 253
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-408-765A-253

Query Match          99.8%; Score 1359; DB 16; Length 267;
Best Local Similarity 99.6%; Pred. No. 1.9e-92;
Matches 266; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKAAVLTTLAVLFLTGSQARHFQQDEPPQSPMDRVKDLATVYVDVLKDSGRDYVSQFEGS 60
Db 1 MKAAVLTTLAVLFLTGSQARHFQQDEPPQSPMDRVKDLATVYVDVLKDSGRDYVSQFEGS 60
QY 61 ALGKQLNLKLDNWDVSTFTSKLRQLGPVTQEFWDNLEKETEGLRQEMSKDLBEVKAK 120
Db 61 ALGKQLNLKLDNWDVSTFTSKLRQLGPVTQEFWDNLEKETEGLRQEMSKDLBEVKAK 120
QY 121 VQPYLDDFQKKWQEEEMELYRQKVEPLRAELQEGARQKLHELOEKLSPGLGEEMRDRARAHV 180
Db 121 VQPYLDDFQKKWQEEEMELYRQKVEPLRAELQEGARQKLHELOEKLSPGLGEEMRDRARAHV 180
QY 181 DALRTHLAPYSDELQRLAARLEALKENGARLAELYHAKATEHLSTLSEKAKPALEDLRQ 240
Db 181 DALRTHLAPYSDELQRLAARLEALKENGARLAELYHAKATEHLSTLSEKAKPALEDLRQ 240
QY 241 GLLPVLESFKVSFLSALEYTKKLNQ 267
Db 241 GLLPVLESFKVSFLSALEYTKKLNQ 267

RESULT 12
US-09-919-039-27
; Sequence 27, Application US/09919039
; Publication No. US20030108871A1
; GENERAL INFORMATION:
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
; FILE REFERENCE: PA-0035 US
; CURRENT APPLICATION NUMBER: US/09/919,039
; CURRENT FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 60/222,113
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 401
; SOFTWARE: PERL Program
; SEQ ID NO 27
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030108871A1 2516070CD1
; US-09-919-039-27

Query Match          98.8%; Score 1346; DB 10; Length 267;
Best Local Similarity 98.9%; Pred. No. 1.8e-91;
Matches 264; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKAAVLTTLAVLFLTGSQARHFQQDEPPQSPMDRVKDLATVYVDVLKDSGRDYVSQFEGS 60
Db 1 MKAAVLTTLAVLFLTGSQARHFQQDEPPQSPMDRVKDLATVYVDVLKDSGRDYVSQFEGS 60
QY 61 ALGKQLNLKLDNWDVSTFTSKLRQLGPVTQEFWDNLEKETEGLRQEMSKDLBEVKAK 120
Db 61 ALGKQLNLKLDNWDVSTFTSKLRQLGPVTQEFWDNLEKETEGLRQEMSKDLBEVKAK 120
QY 121 VQPYLDDFQKKWQEEEMELYRQKVEPLRAELQEGARQKLHELOEKLSPGLGEEMRDRARAHV 180
Db 121 VQPYLDDFQKKWQEEEMELYRQKVEPLRAELQEGARQKLHELOEKLSPGLGEEMRDRARAHV 180
```

Db	121	VQPYLDDFQKKQWEMLYRQVPELRAPLEQGARQKHLHELQKSLPLGEMEDRARAHV	180
Qy	181	DALRTHLAPYSDELQRLAARLEALKKENGSGARLAETYHAKATEHLSTLSEKAKPALEDLRQ	240
Db	181	ARVTHLAPYSDELQRLAARLEALKKENGSGARLAETYHAKATEHLSTLSEKAKPALEDLRQ	240
Qy	241	GLLPVLESFKVSFLSALBEYTKKLNQ	267
Db	241	GLLPVLESFKVSFLSALBEYTKKLNQ	267

RESULT 13

```

US-09-987-107-16
; Sequence 16, Application US/09987107
; Patent No. US20020156007A1
; GENERAL INFORMATION:
; APPLICANT: GRAVERSEN, Jonas
; APPLICANT: MOESTRUP, Soren
; TITLE OF INVENTION: APOLIPOPROTEINS ANALOGUES
; FILE REFERENCE: GRAVERSENIA
; CURRENT APPLICATION NUMBER: US/09/987,107
; CURRENT FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/264,022
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: DK PA2001 00057
; PRIOR FILING DATE: 2001-01-15
; PRIOR APPLICATION NUMBER: DK PA2000 01682
; PRIOR FILING DATE: 2000-11-10
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Macaca fascicularis
US-09-987-107-16

```

Query Match 95.9%; Score 1306; DB 9; Length 267;
Best Local Similarity 95.1%; Pred. No. 1.6e-88;
Matches 254; Conservative 6; Mismatches 7; Indels

[illegible]

RESULT 14

US-10-032-199-77
Sequence 77, Application US/10032189
Publication No. US20030170630A1
GENERAL INFORMATION:
APPLICANT: Alsbrook II, John P
APPLICANT: Tchernev, Velizar T
APPLICANT: Liu, Xiaohong
APPLICANT: Spytek, Kimberly A
APPLICANT: Zerkusen, Bryan D
APPLICANT: Patturajan, Meera
APPLICANT: Grosse, William M

```

1  APPLICANT: Lepley, Denise M
2  APPLICANT: Burgess, Catherine E
3  APPLICANT: Shmkets, Richard A
4  APPLICANT: Grosse, William M
5  APPLICANT: Szekeres, Edward S
6  APPLICANT: Vernet, Corine A.M.
7  APPLICANT: Li, Li
8  APPLICANT: Casman, Stacie J
9  APPLICANT: Boldog, Ferenc L
10 APPLICANT: Gorman, Linda
11 APPLICANT: Gangolli, Esha A
12 APPLICANT: Fernandes, Elma R
13 APPLICANT: Rieger, Daniel K
14 APPLICANT: Edinger, Shlomit R
15 APPLICANT: Gunther, Erik
16 APPLICANT: Millet, Isabelle
17 APPLICANT: Sciore, Paul
18 APPLICANT: Ellerman, Karen
19 APPLICANT: MacDougall, John R
20 APPLICANT: Smatheon, Glennda
21 TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
22 FILE REFERENCE: 21402-228
23 CURRENT APPLICATION NUMBER: US/10/032,189
24 CURRENT FILING DATE: 2001-12-21
25 PRIOR APPLICATION NUMBER: 60/257,495
26 PRIOR FILING DATE: 2000-12-21
27 PRIOR APPLICATION NUMBER: 60/258,171
28 PRIOR FILING DATE: 2000-12-20
29 PRIOR APPLICATION NUMBER: 60/269,940
30 PRIOR FILING DATE: 2001-02-20
31 PRIOR APPLICATION NUMBER: 60/274,192
32 PRIOR FILING DATE: 2001-03-08
33 PRIOR APPLICATION NUMBER: 60/277,826
34 PRIOR FILING DATE: 2001-03-22
35 PRIOR APPLICATION NUMBER: 60/279,840
36 PRIOR FILING DATE: 2001-03-29
37 PRIOR APPLICATION NUMBER: 60/282,981
38 PRIOR FILING DATE: 2001-04-11
39 PRIOR APPLICATION NUMBER: 60/283,656
40 PRIOR FILING DATE: 2001-04-13
41 PRIOR APPLICATION NUMBER: 60/309,247
42 PRIOR FILING DATE: 2001-07-31
43 PRIOR APPLICATION NUMBER: 60/311,754
44 PRIOR FILING DATE: 2001-08-17
45 PRIOR APPLICATION NUMBER: 60/313,331
46 PRIOR FILING DATE: 2001-08-17
47 NUMBER OF SEQ ID NOS: 260
48 SOFTWARE: PatentIn Ver. 2.1
49 SEQ ID NO 77
50 LENGTH: 267
51 TYPE: PRT
52 ORGANISM: Macaca fascicularis
53 US-10-032-189-77

```

Query Match 95.9%: Score 1306: DB 14: Length 267:

Query Match 95.5%; Score 1306; DB 14
Best Local Similarity 95.1%; Pred. No. 1-6e-88;

BEST LOCAL SIMILARITY 93.14; PRED. NO. 1.0E+06;
Matches 254: Conservative 6: Mismatches 7: Indels 0: Gaps 0:

[illegible]

Qy 61 ALGKQLNLKLLDNWDSVTSTFSKLREQLGPVTQEFWDNLKETEGLRQEMSKDLEEVKAK 120

Db 61 ALGKQLNLKLLDNWDSVTSTVSKLREQLGPVTQEFWDNLKETEGLRQEMSKDLEEVKAK 120

Qy	121 VQPYLDDFQKKWOEMELRYQKVEPLRAELQEGARQLHELQEKLSPLGEWMRDRAHV 180 :
D_b	121 VQPYLDDFQKKWOEMELRYQKVEPLRAELHEGTRQKLHELHKEKLSPLGEVMRDRAHV 180 :

[illegible]

QY 241 GLLPVLESFKVSFLSALAEYTKKLTQ 267
Db 241 GLLPVLESFKVSFLSALAEYTKKLTQ 267

RESULT 15

US-10-038-854-223
; Sequence 223, Application US/10038854
; Publication No. US20040022781A1
; GENERAL INFORMATION:
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Li, Li
; APPLICANT: Wolenc, Adam R
; APPLICANT: Vernet, Corine
; APPLICANT: Eisen, Andrew J
; APPLICANT: Liu, Xiaohong
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Shimkets, Richard A
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spaderna, Steven K
; APPLICANT: Gorman, Linda
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Gangolli, Esha A
; APPLICANT: Guo, Xiaojia S
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Rastelli, Luca
; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc
; APPLICANT: Burgess, Catherine E
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gunther, Erik
; APPLICANT: Smithson, Glennda
; APPLICANT: Millet, Isabelle
; APPLICANT: MacDougall, John R
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-230
; CURRENT APPLICATION NUMBER: US/10/038,854
; CURRENT FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/258,928
; PRIOR FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: 60/259,415
; PRIOR FILING DATE: 2001-01-02
; PRIOR APPLICATION NUMBER: 60/259,785
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/269,814
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/279,832
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,833
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,863
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/283,889
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,447
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/286,683
; PRIOR FILING DATE: 2001-04-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 223
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Cynomolgus monkey
US-10-038-854-223

Query Match 95.9%; Score 1306; DB 15; Length 267;
Best Local Similarity 95.1%; Pred. No. 1.6e-88;

Matches 254; Conservative 6; Mismatches 7; Indels 0; Gaps 0;
QY 1 MKAAVLTAVLFLTGSQARHEWQDEPPQSPWDRVKDLATVYVDVKDSGRDYVSQPEGS 60
Db 1 MKATVLTAVLFLTGSQARHFQDEPPQTWDRVKDLVTVYVEALKDSGGDYVSQPEGS 60
QY 61 ALGKQLNLKLLDNWDSVTSTFSLKREQLGPVTQBFWDNLEKETEGLRQEMSKDLEEVKAK 120
Db 61 ALGKQLNLKLLDNWDSVTSTVSKLREQLGPVTQBFWDNLEKETEGLRQEMSKDLEEVKAK 120
QY 121 VQPYLDDFQKKWQEMELYRQKVBPFLRAELOEGARQKLHELOEKLSPILGEEMRDRAHVV 180
Db 121 VQPYLDDFQKKWQEMELYRQKVBPFLRAELHEGTRQKLHELHEKLSPLGEEVDRARAHHV 180
QY 181 DALRTHLAPYSDELQRLAARLEALKENGGAARLAAYHAKATEHLSTLSEKAKPALEDLRQ 240
Db 181 DALRTHLAPYSDELQRLAARLEALKENGGAARLAAYHAKASEHJLSTLSEKAKPALEDLRQ 240
QY 241 GLLPVLESFKVSFLSALAEYTKKLTQ 267
Db 241 GLLPVLESFKVSFLSALAEYTKKLTQ 267

Search completed: December 21, 2004, 07:30:10
Job time : 133.2 secs

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Result No.	Score	Query			DB	ID	Description
		Match	Length				
1	370	41.8	164	2	S21830	apolipoprotein A-I	
2	106	12.0	34	2	S67972	apolipoprotein AI	
3	106	12.0	36	2	A56866	apolipoprotein A-I	
4	101	11.4	20	2	A05313	apolipoprotein A-I	
5	98	11.1	118	2	B69818	conserved hypothet	
6	90.5	10.2	164	2	T32349	hypothetical prote	
7	90	10.2	146	2	G90201	conserved hypothet	
8	88	9.9	129	2	C95164	conserved hypothet	
9	88	9.9	129	2	B98030	conserved hypothet	
10	86.5	9.8	99	2	B46598	ski-related protei	
11	85	9.6	161	2	S48396	tropomyosin TPM2 -	
12	82.5	9.3	163	2	F75559	hypothetical prote	
13	79.5	9.0	135	2	S74746	hypothetical prote	
14	79	8.9	154	2	T12720	hypothetical prote	
15	79	8.9	159	2	A97836	hypothetical prote	
16	78.5	8.9	129	2	F64117	embryonic abundan	
17	78.5	8.9	141	2	C64576	trichohyalin compo	
18	78	8.8	123	2	A84228	hypothetical prote	
19	77	8.7	113	2	S66805	hypothetical prote	
20	77	8.7	142	2	I50496	light meromyosin -	
21	77	8.7	147	2	S70457	plasmaenogen-bindin	
22	77	8.7	148	2	T05610	hypothetical prote	
23	76.5	8.6	79	2	G86722	hypothetical prote	
24	76.5	8.6	162	2	B27475	gas-vesicle protei	
25	76	8.6	109	2	A83265	conserved hypothet	
26	76	8.6	118	2	AD1568	hypothetical prote	
27	76	8.6	158	2	H83434	translocation prot	
28	76	8.6	160	2	H75062	flagella-related p	
29	75.5	8.5	140	2	G72208	hypothetical prote	

Query Match 12.0%; Score 106; DB 2; Length 34;
Best Local Similarity 54.5%; Pred. No. 0.61;
Matches 18; Conservative 9; Mismatches 6; Indels 0; Gaps 0;

QY 2 EPPQSPWDRVKDLATVYVDLVKDSGRDYYVSQFE 34
DB 1 DEQPAPLRLKDLVDVYLETVRASGKDYLAQFE 33

RESULT 3
A56866
apolipoprotein A-I - Japanese quail (fragment)
C:Species: Coturnix coturnix japonica (Japanese quail)
C>Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 09-Jul-2004
C:Accession: A56866
R:Oku, H.; Ishikawa, M.; Nagata, J.; Toda, T.; Chinen, I.
Biochim. Biophys. Acta 1167, 22-28, 1993
A:Title: Lipoprotein and apoprotein profile of Japanese quail.
A:Reference number: A56866; MUID:93213845; PMID:8461329
A:Accession: A56866
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-36 <OKU>
A:CROSS-references: UNIPROT:P32918
A:Note: sequence extracted from NCBI backbone (NCBIP:128831)
C:Superfamily: apolipoprotein A-I

Query Match 12.0%; Score 106; DB 2; Length 36;
Best Local Similarity 50.0%; Pred. No. 0.65;
Matches 18; Conservative 10; Mismatches 8; Indels 0; Gaps 0;

QY 2 EPPQSPWDRVKDLATVYVDLVKDSGRDYYVSQFESSA 37
DB 1 DDPQPLDIRDMLVDVYLETVRASGKDAISQFESSA 36

RESULT 4
A05313
apolipoprotein A-I - red guenon (fragment)
N:Alternate names: apo-A-I
C:Species: Erythrocebus patas (red guenon, hussar)
C>Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 09-Jul-2004
C:Accession: A05313
R:Mahley, R.W.; Weisgraber, K.H.; Innerarity, T.; Brewer Jr., H.B.
Biochemistry 15, 1928-1933, 1976
A:Title: Characterization of the plasma lipoproteins and apoproteins of the Erythrocebus
A:Reference number: A90395; MUID:76184721; PMID:178359
A:Accession: A05313
A:Molecule type: protein
A:Residues: 1-20 <MAH>
A:CROSS-references: UNIPROT:P18647
C:Superfamily: apolipoprotein A-I
C:Keywords: cholesterol metabolism; HDL; intestine; lipid binding; lipid transport; lipid

Query Match 11.4%; Score 101; DB 2; Length 20;
Best Local Similarity 85.0%; Pred. No. 0.71;
Matches 17; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 DEPPQSPWDRVKDLATVYVD 20
DB 1 DEPPQTPWDRVKDLVTYYVE 20

RESULT 5
B69818
conserved hypothetical protein yhaH - Bacillus subtilis
C:Species: Bacillus subtilis
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C:Accession: B69818
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte
C.; Bron, S.; Brouillet, S.; Prutei, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,

A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpatra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: B69818
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-118 <KUN>
A:CROSS-references: UNIPROT:O07516; GB:Z99109; GB:AL009126; NID:92633260; PIDN:CAB12840
A:Experimental source: strain 168
C:Genetics:
A:Gene: yhaH

Query Match 11.1%; Score 98; DB 2; Length 118;
Best Local Similarity 27.5%; Pred. No. 7.6;
Matches 28; Conservative 25; Mismatches 33; Indels 16; Gaps 3;

QY 39 GKQLNLKLLDNWDSVTSTFSKLRQLGPVTOEFPWNLEKETEGLR--OEMSKDLEEVKAK 96
DB 30 GKQLREKMTNYDSFEETIKELKS-----DGLALKDQLIKAAKESDVIKDVGGE 79

QY 97 VQPYLDDPQKKQWEMELYRQKVEPLRAELOEGARQKLHELO 138
DB 80 LQTSI---KKWBEIKPHQODLQKEIADIEKIRLEKTLQ 117

RESULT 6
T23249
hypothetical protein K02E2.7 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T23249
R:Lloyd, C.
Submitted to the EMBL Data Library, November 1996
A:Reference number: Z19716
A:Accession: T23249
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-164 <WIL>
A:CROSS-references: UNIPROT:Q9XUV5; EMBL:Z81560; PIDN:CAB04544.1; GSPDB:GN00023; CESP:K02E2
A:Experimental source: clone K02E2
C:Genetics:
A:Gene: CESP:K02E2.7
A:Map position: 5
A:Introns: 35/1; 62/1

Query Match 10.2%; Score 90.5; DB 2; Length 164;
Best Local Similarity 22.7%; Pred. No. 32;
Matches 40; Conservative 27; Mismatches 52; Indels 57; Gaps 7;

QY 2 EP---PQSPWDRVKDLATVYVDLVKDSGRDYYVSQFEGSALGK-----QLNLKLL 47
DB 9 EPELQPTSPPPSIHD-----DLINDMDGGMFGSEKTTDSSLEDLQJHVQEM 54

QY 48 DNWDSVTSTFSKLRQLGPVTOEFPWN-----LEK-----ETEGLRQEMSK 88
DB 55 DKNTMLTENFDQDQELKLFVQELNQAAETINSSHQLENALKAAENIETISEEASD 114

QY 89 DLEEVKAKVPLYDDPQKKQWEMELYRQKVEPLRAELOEGARQKLHELOKLSPL 144
DB 115 DEEE-----DNFHTMLQEQAAARHKEEMKVMKKRAE-QKKKLEELNRRKNVAI 160

RESULT 7
G90201
conserved hypothetical protein [imported] - Sulfolobus solfataricus
C:Species: Sulfolobus solfataricus
C>Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004

C:Accession: G90201
R:She, Q.; Jing, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, H.
aret, R.A.; Ragan, M.A.; Senses, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A:Description: Sulfolobus solfataricus complete genome.
A:Reference number: A99139
A:Accession: G90201
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-146 <KUR>
A:Cross-references: UNIPROT:Q9UW7; GB:AE006641; NID:gl3813718; PIDN:AAK40870.1; GSPDB:G
C:Genetics:
A:Gene: SSO553

Query Match 10.2%; Score 90; DB 2; Length 146;
Best Local Similarity 22.4%; Pred. No. 30;
Matches 33; Conservative 34; Mismatches 54; Indels 26; Gaps 5;
QY 21 VLKDSGR---DYVSQFEGSALGKQLNLKLLDNWDSVTSTFKLREQLG-----PVTQEPF 72
DB 6 VKSNGKVTSEKVAEGEISKVLRDVAKAEALSWNELASDFIIMRDNQEVRLPLPKPEVY 65
QY 73 DNLEKETELQEMSKDLEEVKAKVQPYLDDFKKKQWE-----EMELY-----RQKVE 120
DB 66 EAIKFLIG-----KDKKEAIKIPVYIISYENWESDFQDKKIYVVSFYINDEIKD 119
QY 121 PLRAELQEGARQKLHLEQKLSPLGEE 147
DB 120 VLNDAAQMTSEQKQBELKEELBEEBEE 146

RESULT 8
C95164
conserved hypothetical protein SP1410 [imported] - Streptococcus pneumoniae (strain TIGR
C:Species: Streptococcus pneumoniae
C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
C:Accession: C95164
R:Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
on, J.D.; Unayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A:Reference number: A95000; MUID:21357209; PMID:11463916
A:Accession: C95164
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-129 <KUR>
A:Cross-references: UNIPROT:Q97Q21; GB:AE005672; PIDN:AAK75508.1; PID:gl4972899; GSPDB:G
A:Experimental source: strain TIGR4
C:Genetics:
A:Gene: SP1410

Query Match 9.9%; Score 88; DB 2; Length 129;
Best Local Similarity 25.8%; Pred. No. 36;
Matches 25; Conservative 27; Mismatches 41; Indels 4; Gaps 1;
QY 62 EQLGPVTQEFWNLEKETELQEMSKDLEEVKAKVQPYLDDFKKKQWEMELYRQKVEP 121
DB 27 KQVCSQAQDFLDDLRDPYAKEQVCEKLTVEKEQA---TDFVLKTKQVSGEITVDS 82
QY 122 LRAELQEGARQKLHLEQKLSPLGEMDRARAHVDA 158
DB 83 ILAQTYSYAFQATEASKQNLLNLEQWQEKAEALDSD 119

RESULT 9
B98030
conserved hypothetical protein spr1267 [imported] - Streptococcus pneumoniae (strain R6)
C:Species: Streptococcus pneumoniae
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C:Accession: B98030

R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Bargett, S.; DeHoff, B.S.; t
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; t
Y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A:Reference number: A97872; MUID:21429245; PMID:11544234
A:Accession: B98030
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-129 <KUR>
A:Cross-references: UNIPROT:Q8DPA5; GB:AE007317; PIDN:AAAL00071.1; PID:gl5458907; GSPDB:B
C:Genetics:
A:Gene: spr1267

Query Match 9.9%; Score 88; DB 2; Length 129;
Best Local Similarity 25.8%; Pred. No. 36;
Matches 25; Conservative 27; Mismatches 41; Indels 4; Gaps 1;
QY 62 EQLGPVTQEFWNLEKETELQEMSKDLEEVKAKVQPYLDDFKKKQWEMELYRQKVEP 121
DB 27 KQVCSQAQDFLDDLRDPYAKEQVCEKLTVEKEQA---TDFVLKTKQVSGEITVDS 82
QY 122 LRAELQEGARQKLHLEQKLSPLGEMDRARAHVDA 158
DB 83 ILAQAQSYAFQATEASKQNLLNLEQWQEKAEALDSD 119

RESULT 10
B46598
ski-related protein Snon - chicken (fragments)
C:Species: Gallus gallus (chicken)
C:Date: 03-Mar-1994 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C:Accession: B46598
R:Nagase, T.; Nomura, N.; Ishii, S.
J. Biol. Chem. 268, 13710-13716, 1993
A:Title: Complex formation between proteins encoded by the ski gene family.
A:Reference number: A46598; MUID:93293901; PMID:8514802
A:Accession: B46598
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-56;57-99 <NAG>
A:Cross-references: UNIPROT:Q9PS83
A:Note: sequence extracted from NCBI backbone (NCBIP:134568, NCBIP:134572)
C:Superfamily: ski transforming protein
C:Keywords: DNA binding; tandem repeat

Query Match 9.8%; Score 86.5; DB 2; Length 99;
Best Local Similarity 30.7%; Pred. No. 33;
Matches 31; Conservative 21; Mismatches 28; Indels 21; Gaps 4;
QY 41 QLNLKLLDNWDSVTSTFKLREQLGPTQEFWNLEKETELQEMSKDLEEVKAKVQPY 100
DB 17 QMEVNLSS-----SKSMKL---TEE-QQNLQKLESLQNEHAQRMESQLAEALR 64
QY 101 LDDFOKKQWEMELYRQKVEPLRAELQEGARQKLHLEQKL 141
DB 65 LDHAEDAQELQDELQERE-----ARQKLEWMIKEL 96

RESULT 11
S48396
tropomyosin TPM2 - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein YIL138c
C:Species: Saccharomyces cerevisiae
C:Date: 02-Dec-1994 #sequence_revision 02-Dec-1994 #text_change 09-Jul-2004
C:Accession: S48396; A56490
R:Churcher, C.
submitted to the EMBL Data Library, September 1994
A:Reference number: S48310
A:Accession: S48396
A:Molecule type: DNA
A:Residues: 1-161 <CHU>

A:Accession: A97836
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-159 <KUR>
A:Cross-references: UNIPROT:Q92GN4; GB:AE006914; PIDN:AAL03627.1; PID:g15620212; GSPDB:C
C:Genetics:
A:Gene: RC1089

Query Match 8.9%; Score 79; DB 2; Length 159;
Best Local Similarity 24.4%; Pred. No. 1.6e+02;
Matches 19; Conservative 17; Mismatches 42; Indels 0; Gaps 0;
QY 51 DSVTSTFSKLRQLGPPVTOEFWDNLEKETEGLRQEMSKOLEEVKAKVQPYLDDFQKKWQE 110
Db 72 DRFNQTLSTYDQIAQSVTEIKKNVQDKDKKGEELQKSIDAVKAKWEAYKKAGSKQEK 131
QY 111 EMELYRQKVEPLRAELQE 128
Db 132 IRQHLVDKLEELNKNINE 149

Search completed: December 21, 2004, 10:52:52
Job time : 18 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2004, 13:34:52 ; Search time 187 Seconds
(without alignments)
523.068 Million cell updates/sec

Title: US-09-803-918A-2_COPY_25_194

Perfect score: 885
Sequence: 1 DEPPQSPMDRVKDLATVYVD.....RAAHVDALRTHLAPYSDEL 170

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 626041

Minimum DB seq length: 0
Maximum DB seq length: 170

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_02: *
1: uniprot_sprot: *
2: uniprot_trembl: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	320	36.2	79	2	Q6LD50	Q6ld50 mus sp. apo
2	320	36.2	79	2	AAB35539	Aab35539 mus sp. a
3	247	27.9	56	2	O02762	O02762 ovis aries
4	217	24.5	82	2	Q29248	Q29248 sus scrofa
5	154	17.4	67	2	Q9Y355	Q9y355 homo sapien
6	141	15.9	26	2	Q9UCT8	Q9uct8 homo sapien
7	113	12.8	155	2	Q6CMT0	Q6cmt0 kluyveromyc
8	111	12.5	150	2	Q9XDR3	Q9xdr3 streptococc
9	110.5	12.5	138	2	Q9S613	Q9s613 streptococc
10	106	12.0	34	2	Q9PRR6	Q9pr6 anser anser
11	103	11.6	143	1	LEA1 APHAV	Q9Sv77 aphelenchus
12	101.5	11.5	107	1	APE MACMU	Q28502 macaca mula
13	101	11.4	20	1	APAI1 ERYPA	P18647 erythrocebu
14	101	11.4	30	2	Q9QV04	Q9qv04 rattus sp.
15	99	11.2	107	1	APE SAISC	Q28995 saimiri sci
16	98.5	11.1	129	2	Q8ZX78	Q8zx78 pyrobaculum
17	98	11.1	117	2	Q6QZ11	Q6qzil pseudopleur
18	98	11.1	117	2	AAT01918	Aat01918 pseudople
19	98	11.1	118	2	O07516	O07516 bacillus su
20	97	11.0	96	2	Q8RH51	Q8rh51 fusobacteri
21	95.5	10.8	163	2	Q9R0H4	Q9r0h4 rattus norv
22	94	10.6	166	2	Q63297	Q63297 rattus norv
23	93.5	10.6	163	2	Q9ZF11	Q9zfl1 streptococc
24	92.5	10.5	102	2	Q98TJ9	Q98tj9 platicithys
25	92.5	10.5	148	2	Q84MC4	Q84mc4 arabidopsis
26	92	10.4	114	2	Q6T257	Q6t257 crassostrea
27	92	10.4	114	2	AAR17060	Aar17060 crassostr
28	92	10.4	168	2	Q9AGA2	Q9aga2 streptococc
29	90.5	10.2	164	2	Q9XUV5	Q9xuv5 caenorhabdi
30	90	10.2	146	2	Q9UWV7	Q9uwv7 sulfolobus
31	89	10.1	168	2	Q9SR63	Q9sr63 arabidopsis

32	88.5	10.0	161	2	Q8FKV5	Q8fkv5 candida gla
33	88	9.9	96	2	Q7P280	Q7p280 fusobacteri
34	88	9.9	121	2	Q6YQ37	Q6yq37 onion yello
35	88	9.9	121	2	BAD04623	Bad04623 onion yel
36	88	9.9	122	1	YSB6 METAC	O8tin0 methanosarc
37	88	9.9	129	2	Q8DPA5	Q8dpa5 streptococc
38	88	9.9	129	2	Q97Q21	Q97q21 streptococc
39	87.5	9.9	170	2	Q74MC1	Q74mc1 nanoarchaeu
40	87.5	9.9	170	2	AAR39327	Aar39327 nanoarcha
41	87	9.8	118	2	O68959	O68959 streptococc
42	87	9.8	158	2	Q9RN96	Q9rn96 streptococc
43	87	9.8	170	2	Q72JB9	Q72jb9 thermus the
44	87	9.8	170	2	AAS81203	Aas81203 thermus t
45	86.5	9.8	99	2	Q9PSS3	Q9pes3 gallus gall

ALIGNMENTS

RESULT 1						
Q6LD50						
ID	Q6LD50	PRELIMINARY;	PRT;	79	AA.	
AC	Q6LD50;					
DT	05-JUL-2004 (TremBLrel. 27, Created)					
DT	05-JUL-2004 (TremBLrel. 27, Last sequence update)					
DT	05-JUL-2004 (TremBLrel. 27, Last annotation update)					
DE	Apolipoprotein A1 homolog protein (Fragment).					
GN	Name=apolipoprotein A1 homolog;					
OS	Mus sp.					
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.					
OX	NCBI_TaxID=10095;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RX	MEDLINE=96062440; PubMed=7488287;					
RA	Merrill J.F., Rivkin E., Shen C., Lahita R.G.;					
RT	"Selection of a gene for apolipoprotein A1 using autoantibodies from a					
RT	patient with systemic lupus erythematosus.";					
RL	Arthritis Rheum. 38:1655-1659(1995).					
DR	EMBL; S80442; AAB35539.1; -.					
DR	InterPro; IPR009074; Apolipo_A_E_C3.					
KW	Lipoprotein.					
FT	NON_TER	79				
SQ	SEQUENCE	79	AA;	9583	MW;	33CA72DA854A150A CRC64;
Query Match						
Best Local Similarity 36.2%; Score 320; DB 2; Length 79;						
Matches 60; Conservative 10; Mismatches 8; Indels 0; Gaps 0;						
QY 70 BFWNDLKEKETGLRQEMSKLLEEVKAKVQPYLDDFQKKWQEMELYRQKVEPLRAELQSG 129						
Db 1 DFWDNLKXETDWRQEMNKDLEEVKQKQPYLDFQKKWEDVELYRQKVAFLGAELQSS 60						
QY 130 ARQKLHLEQLKSLPGEE 147						
Db 61 ARQKLQELQGRLLSPVAEE 78						
RESULT 2						
AAB35539						
ID	AAB35539	PRELIMINARY;	PRT;	79	AA.	
AC	AAB35539;					
DT	02-MAR-2004 (TremBLrel. 27, Created)					
DT	02-MAR-2004 (TremBLrel. 27, Last sequence update)					
DT	02-MAR-2004 (TremBLrel. 27, Last annotation update)					
DE	Apolipoprotein A1 homolog (Fragment).					
OS	Mus sp.					
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.					
OX	NCBI_TaxID=10095;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RX	MEDLINE=96062440; PubMed=7488287;					

```

RA Merrill J.T., Rivkin E., Shen C., Lahita R.G.;
RT "Selection of a gene for apolipoprotein A1 using autoantibodies from a
RL patient with systemic lupus erythematosus.";
RL Arthritis Rheum. 38:1655-1659(1995).
FT EMBL; S80442; AAB35539.1; -.
FT NON_TER 1
SQ SEQUENCE 79 AA; 9583 MW; 33CA72DA854A150A CRC64;

Query Match 36.2%; Score 320; DB 2; Length 79;
Best Local Similarity 76.9%; Pred. No. 3.9e-13;
Matches 60; Conservative 10; Mismatches 8; Indels 0; Gaps 0;

QY 70 EFWNDLEKTEGLRQEMSKDLBEVKAKQVPYLDDFKKQWQEMELYRKVEFLRAELQEG 129
:|||||: :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 DFWNDLEKETDWRQEMNKDLBEVKAKQVPYLDEFKKWKEDEVELYRKQVAPLGAELQES 60

QY 130 ARQKLHELQELSPGEE 147
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 ARQKLQELQRLSPVAEE 78

RESULT 3
O02762 PRELIMINARY; PRT; 56 AA.
AC O02762;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Apolipoprotein A1 (Fragment).
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99098454; PubMed=9883985;
RA Robertson J.A., Bhattacharyya S., Ing N.H.;
RT "Tamoxifen up-regulates oestrogen receptor-alpha, c-fos and
RT glyceraldhyde 3-phosphate-dehydrogenase mRNAs in ovine endometrium.";
RL J. Steroid Biochem. Mol. Biol. 67:285-292(1998).
DR EMBL; U94720; AAB57840.1; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0008289; F:lipid binding; IEA.
DR GO; GO:0006869; P:lipid transport; IEA.
DR GO; GO:0042157; P:lipoprotein metabolism; IEA.
DR InterPro; IPR000074; Apolipoprotein.
DR InterPro; IPR009074; Apolipo_A_E_C3.
DR Pfam; PF01442; Apolipo_A_E_C3.
KW Lipoprotein.
FT NON_TER 1
FT NON_TER 56
SQ SEQUENCE 56 AA; 6617 MW; 2AB38E08F1B8F1BC CRC64;

Query Match 27.9%; Score 247; DB 2; Length 56;
Best Local Similarity 82.1%; Pred. No. 1e-08;
Matches 46; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 47 LDNWDVSTFTSKLEQGLGPVTQEFWDNLKETEGLRQEMSKDLBEVKAKQVPYLD 102
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 LDNWDVSLSTSKVREQLGPVTQEFWDNLKETASLRQEMHKDLEEARQKQVPYLD 56

RESULT 4
Q29248 PRELIMINARY; PRT; 82 AA.
AC Q29248;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Apolipoprotein A-I (Fragment).
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Small intestine;
RC MEDLINE=96327607; PubMed=8672129;
RA Wintxoe A.K., Fredholm M., Davies W.;
RT "Evaluation and characterization of a porcine small intestine cDNA
RT library.";
RL Mamm. Genome 7:509-517(1996).
DR EMBL; F14858; CAA23298.1; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0008289; F:lipid binding; IEA.
DR GO; GO:0006869; P:lipid transport; IEA.
DR GO; GO:0042157; P:lipoprotein metabolism; IEA.
DR InterPro; IPR000074; Apolipoprotein.
DR InterPro; IPR009074; Apolipo_A_E_C3.
DR Pfam; PF01442; Apolipo_A_E_C3.
KW Lipoprotein.
FT NON_TER 82
FT NON_TER 82
SQ SEQUENCE 82 AA; 9168 MW; 24625C65CBFFEDD8 CRC64;

Query Match 24.5%; Score 217; DB 2; Length 82;
Best Local Similarity 70.7%; Pred. No. 1.2e-06;
Matches 41; Conservative 5; Mismatches 12; Indels 0; Gaps 0;

QY 2 EPPQSPWDRVKDLATVYVDLKDGRDYSQFEGSALGKQLNKLNDWDSVTTFSK 59
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 25 DDPXSPWDRVKDFTXVYVDAIKDSGRDYXAQFEASALGKHLNKLXNDXSLGXTFTK 82

RESULT 5
Q9Y355 PRELIMINARY; PRT; 67 AA.
AC Q9Y355;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Apolipoprotein A1 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99218322; PubMed=10198255;
RA Hamidi Asl K., Liepnieks J.J., Nakamura M., Parker F., Benson M.D.;
RT "A novel apolipoprotein A-1 variant, Arg173Pro, associated with
RT cardiac and cutaneous amyloidosis.";
RL Biochem. Biophys. Res. Commun. 257:584-588(1999).
DR EMBL; AF148963; AAD34604.1; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0008289; F:lipid binding; IEA.
DR GO; GO:0006869; P:lipid transport; IEA.
DR GO; GO:0042157; P:lipoprotein metabolism; IEA.
DR InterPro; IPR000074; Apolipoprotein.
DR InterPro; IPR009074; Apolipo_A_E_C3.
DR Pfam; PF01442; Apolipoprotein; 1.
KW Lipoprotein.
FT NON_TER 1
FT NON_TER 67
SQ SEQUENCE 67 AA; 7433 MW; 525E1FEB7BDD5AFB CRC64;

Query Match 17.4%; Score 154; DB 2; Length 67;
Best Local Similarity 100.0%; Pred. No. 0.0088;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 141 LSPGGEWRDRARAHVDALRTHLAPYSDEL 170
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 LSPGGEWRDRARAHVDALRTHLAPYSDEL 30

RESULT 6

```


RP	SEQUENCE FROM N.A.	RC	STRAIN=NRRL Y-1140;	RA	Genoscope;	RL	Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.	RD	EMBL; CR382125; CAG99846.1; --	RE	NON_TER	RF	NON_TER	SQ	SEQUENCE	155 AA; 18171 MW; 3B9137FD16A4E39A CRC64;	Query Match	12.8%; Score 113; DB 2; Length 155;	Best Local Similarity	30.5%; Pred. No. 8;	Matches	43; Conservative	27; Mismatches	47; Indels	24; Gaps	17	
QY	40	KQINLKL-LDNWDSVTSTFSKLRQLGVPVTOEFWDN-----LEKETEGLRQE	85	1	KLNLKLAEASNQ---DKYEELKEQVKVLEQKTDNENLIKSLTTKQQTLEEEVETLEG	57	86	MS--KDLIEVKAQVPYLLDDFKKKWQ---EMELYRKVKEPLRAELQEGARQKLHELOEK	140	58	VSEHQQLAEDSTNLKSHNENFSKNNQQLKEELEENDSKLTKETEKURE-ADLKVEQLERK	116	141	LSPGLGEEMRDRAAHVDALRT	161	117	VTSL-EQERDDLEKKYDDLQT	136	141	LSPGLGEEMRDRAAHVDALRT	161	117	VTSL-EQERDDLEKKYDDLQT	136	141	LSPGLGEEMRDRAAHVDALRT	161
QY	86	MS--KDLIEVKAQVPYLLDDFKKKWQ---EMELYRKVKEPLRAELQEGARQKLHELOEK	140	58	VSEHQQLAEDSTNLKSHNENFSKNNQQLKEELEENDSKLTKETEKURE-ADLKVEQLERK	116	141	LSPGLGEEMRDRAAHVDALRT	161	117	VTSL-EQERDDLEKKYDDLQT	136	141	LSPGLGEEMRDRAAHVDALRT	161	117	VTSL-EQERDDLEKKYDDLQT	136	141	LSPGLGEEMRDRAAHVDALRT	161	117	VTSL-EQERDDLEKKYDDLQT	136	141	LSPGLGEEMRDRAAHVDALRT	161
QY	141	LSPGLGEEMRDRAAHVDALRT	161	117	VTSL-EQERDDLEKKYDDLQT	136	141	LSPGLGEEMRDRAAHVDALRT	161	117	VTSL-EQERDDLEKKYDDLQT	136	141	LSPGLGEEMRDRAAHVDALRT	161	117	VTSL-EQERDDLEKKYDDLQT	136	141	LSPGLGEEMRDRAAHVDALRT	161	117	VTSL-EQERDDLEKKYDDLQT	136	141	LSPGLGEEMRDRAAHVDALRT	161
QY	117	VTSL-EQERDDLEKKYDDLQT	136	141	LSPGLGEEMRDRAAHVDALRT	161	117	VTSL-EQERDDLEKKYDDLQT	136	141	LSPGLGEEMRDRAAHVDALRT	161	117	VTSL-EQERDDLEKKYDDLQT	136	141	LSPGLGEEMRDRAAHVDALRT	161	117	VTSL-EQERDDLEKKYDDLQT	136	141	LSPGLGEEMRDRAAHVDALRT	161	117	VTSL-EQERDDLEKKYDDLQT	136
QY	141	LSPGLGEEMRDRAAHVDALRT	161	117	VTSL-EQERDDLEKKYDDLQT	136	141	LSPGLGEEMRDRAAHVDALRT	161	117	VTSL-EQERDDLEKKYDDLQT	136	141	LSPGLGEEMRDRAAHVDALRT	161	117	VTSL-EQERDDLEKKYDDLQT	136	141	LSPGLGEEMRDRAAHVDALRT	161	117	VTSL-EQERDDLEKKYDDLQT	136	141	LSPGLGEEMRDRAAHVDALRT	161
QY	117	VTSL-EQERDDLEKKYDDLQT	136	141	LSPGLGEEMRDRAAHVDALRT	161	117	VTSL-EQERDDLEKKYDDLQT	136	141	LSPGLGEEMRDRAAHVDALRT	161	117	VTSL-EQERDDLEKKYDDLQT	136	141	LSPGLGEEMRDRAAHVDALRT	161	117	VTSL-EQERDDLEKKYDDLQT	136	141	LSPGLGEEMRDRAAHVDALRT	161	117	VTSL-EQERDDLEKKYDDLQT	136
QY	141	LSPGLGEEMRDRAAHVDALRT	161	117	VTSL-EQERDDLEKKYDDLQT	136	141	LSPGLGEEMRDRAAHVDALRT	161	117	VTSL-EQERDDLEKKYDDLQT	136	141	LSPGLGEEMRDRAAHVDALRT	161	117	VTSL-EQERDDLEKKYDDLQT	136	141	LSPGLGEEMRDRAAHVDALRT	161	117	VTSL-EQERDDLEKKYDDLQT	136	141	LSPGLGEEMRDRAAHVDALRT	161
QY	117	VTSL-EQERDDLEKKYDDLQT	136	141	LSPGLGEEMRDRAAHVDALRT	161	117	VTSL-EQERDDLEKKYDDLQT	136	141	LSPGLGEEMRDRAAHVDALRT	161	117	VTSL-EQERDDLEKKYDDLQT	136	141	LSPGLGEEMRDRAAHVDALRT	161	117	VTSL-EQERDDLEKKYDDLQT	136	141	LSPGLGEEMRDRAAHVDALRT	161	117	VTSL-EQERDDLEKKYDDLQT	136
QY	141	LSPGLGEEMRDRAAHVDALRT	161	117	VTSL-EQERDDLEKKYDDLQT	136	141	LSPGLGEEMRDRAAHVDALRT	161	117	VTSL-EQERDDLEKKYDDLQT	136	141	LSPGLGEEMRDRAAHVDALRT	161	117	VTSL-EQERDDLEKKYDDLQT	136	141	LSPGLGEEMRDRAAHVDALRT	161	117	VTSL-EQERDDLEKKYDDLQT	136	141	LSPGLGEEMRDRAAHVDALRT	161
QY	117	VTSL-EQERDDLEKKYDDLQT	136	141	LSPGLGEEMRDRAAHVDALRT	161	117	VTSL-EQERDDLEKKYDDLQT	136	141	LSPGLGEEMRDRAAHVDALRT	161	117	VTSL-EQERDDLEKKYDDLQT	136	141	LSPGLGEEMRDRAAHVDALRT	161	117	VTSL-EQERDDLEKKYDDLQT	136	141	LSPGLGEEMRDRAAHVDALRT	161	117	VTSL-EQERDDLEKKYDDLQT	136
QY	141	LSPGLGEEMRDRAAHVDALRT	161	117	VTSL-EQERDDLEKKYDDLQT	136	141	LSPGLGEEMRDRAAHVDALRT	161	117	VTSL-EQERDDLEKKYDDLQT	136	141	LSPGLGEEMRDRAAHVDALRT	161	117	VTSL-EQERDDLEKKYDDLQT	136	141	LSPGLGEEMRDRAAHVDALRT	161	117	VTSL-EQERDDLEKKYDDLQT	136	141	LSPGLGEEMRDRAAHVDALRT	161
QY	117	VTSL-EQERDDLEKKYDDLQT	136	141	LSPGLGEEMRDRAAHVDALRT	161	117	VTSL-EQERDDLEKKYDDLQT	136	141	LSPGLGEEMRDRAAHVDALRT	161	117	VTSL-EQERDDLEKKYDDLQT	136	141	LSPGLGEEMRDRAAHVDALRT	161	117	VTSL-EQERDDLEKKYDDLQT	136	141	LSPGLGEEMRDRAAHVDALRT	161	117	VTSL-EQERDDLEKKYDDLQT	136
QY	141	LSPGLGEEMRDRAAHVDALRT	161	117	VTSL-EQERDDLEKKYDDLQT	136	141	LSPGLGEEMRDRAAHVDALRT																			

```
GN Name=emm;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
ON NCBI_TaxID=1314;
RX STRAIN=4973;
RP SEQUENCE FROM N.A.
RC Beall B.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF128960; RAD28609.1; -.
FT NON_TER 1
FT NON_TER 138
SQ SEQUENCE 138 AA; 15656 MW; A50AE51B14EAAE1P CRC64;

Query Match 12.5%; Score 110.5; DB 2; Length 138;
Best Local Similarity 25.6%; Pred. No. 10;
Matches 34; Conservative 20; Mismatches 54; Indels 25; Gaps 2;

QY 35 GSALGKQLNLKLLDWDSDVTSTFSKLREQLGPVTQEFWDNLEKETEGLEQEMS----- 87
Db 11 GAGLASQTEVKAEDHWRVKNARNQVLGRSTVTPAHFWENQRREIKLGEIDQLKLLG 70
QY 88 -----KDL EEVKAKVQVPLDQFQKWOEEMELYRQKVEPLRAELQEGAROKL 134
Db 71 KSNFNDNDLKERLKDLSDEKQIQELEKNKQKDYQLAVDGHANTVVRHEAEI-----AKL 125
QY 135 HELQKLSPLGEE 147
Db 126 DEIIKKQSETABE 138

RESULT 10
Q9PRR6 PRELIMINARY; PRT; 34 AA.
AC Q9PRR6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Apolipoprotein AI (Fragment).
OS Anser anser (domestic goose).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Anseriformes; Anatidae; Anser.
OX NCBI_TaxID=8843;
RN [1];
RP SEQUENCE.
RX MEDLINE=96128192; PubMed=8536707;
RA Hermier D., Sellier N., Rousselot-Pailley D., Forgez P.;
RT "Characterization of apolipoproteins B-100, AI and C from plasma
RT lipoprotein in the goose, Anser anser. Evidence for a genetic
RT polymorphism in Apoc-Like apolipoproteins.",
RL Eur. J. Biochem. 234:586-591(1995).
DR PIR; S67972; S67972.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0008289; F:lipid binding; IEA.
DR GO; GO:0006869; P:lipid transport; IEA.
DR GO; GO:0042157; P:lipoprotein metabolism; IEA.
DR InterPro; IPR000074; P:lipoprotein.
DR InterPro; IPR009074; Apolipo_A_E_C.
DR Pfam; PF01442; Apolipoprotein; 1.
SQ SEQUENCE 34 AA; 3838 MW; 4BE7DFA02BF1DE91 CRC64;

Query Match 12.0%; Score 106; DB 2; Length 34;
Best Local Similarity 54.5%; Pred. No. 4.5;
Matches 18; Conservative 9; Mismatches 6; Indels 0; Gaps 0;

QY 2 BPPQSPWRDRKDLATVYVDVLKDSGRDYVSQPE 34
Db 1 DEFPAPLRKDLVDVYLTVKASGNDYLAQPE 33

RESULT 11
LEA1_APHAV
```

```
ID AC Q9SV77; PRT; 143 AA.
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 29-MAR-2004 (Rel. 43, Last annotation update)
DE Late embryogenesis abundant protein 1 (Aavleal).
OS Aphelenchus avenae (Mycophagous nematode).
OC Eukaryota; Metazoa; Chromadorea; Tylenchida; Aphelenchoidea;
OC Aphelenchidae; Aphelenchus.
OX NCBI_TaxID=70226;
RN [1];
RP SEQUENCE FROM N.A.; AND INDUCTION.
RX MEDLINE=21877401; PubMed=11882885; DOI=10.1038/416038a;
RA Browne J., Tunnacliffe A., Burrell A.;
RA "Anydrosiosis: plant desiccation gene found in a nematode.";
RL Nature 416:38-38(2002).
CC -!- FUNCTION: May be involved in defense against water stress.
CC -!- INDUCTION: Up-regulated in response to desiccation stress.
CC -!- SIMILARITY: Belongs to the LEA type 1 family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; AF423069; AAL18843.1; -.
DR InterPro; IPR004238; LEA.
DR Pfam; PF02987; LEA_4; 2.
KW Repeat.
FT DOMAIN 47 101 4 X 11 AA APPROXIMATE REPEATS.
FT REPEAT 47 57 1.
FT REPEAT 69 79 2.
FT REPEAT 80 90 3.
FT REPEAT 91 101 4.
SQ SEQUENCE 143 AA; 16028 MW; A41566258765BFB0 CRC64;

Query Match 11.6%; Score 103; DB 1; Length 143;
Best Local Similarity 24.1%; Pred. No. 31;
Matches 26; Conservative 27; Mismatches 51; Indels 4; Gaps 1;

QY 45 KLLDNWSDVTSTFSKLREQLGPVTQEFWDNLEKETEGLEQEMS KDL EEVKAKVQVPLDQF 104
Db 25 KVNVAWESTKETLSSTAQAAAEKTAEPDSDAGETIRDLTGQAQKQSGFKRAGEKABET 84
QY 105 QKKWOEEMELYRQKVEPLRAELQEGAROKLHELQKLSPLGEMRDRA 152
Db 85 KQAGEKMDETKQRA----GEMRENAGQKMEYKQOGKGAELRDTA 128

RESULT 12
APE_MACMU STANDARD; PRT; 107 AA.
ID APE MACMU
AC Q28502;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Apolipoprotein E (Apo-E) (Fragment).
GN Name=APOE;
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9544;
RN [1];
RP SEQUENCE FROM N.A.
RX MEDLINE=96225955; PubMed=8635577;
RA Morelli L., Wei L., Amorim A., McDermid J., Abec C.R., Frangione B.,
RA Walker L.C., Levy E.;
RT "Cerebrovascular amyloidosis in squirrel monkeys and rhesus monkeys:
RT apolipoprotein E genotype.";
```

```

CC cholesterol acyltransferase (LCAT).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Major protein of plasma HDL, also found in
CC chylomicrons.
CC -1- SIMILARITY: Belongs to the apolipoprotein A1/A4/E family.
DR PIR: A05313; A05313; Apolipo_A_E_C3.
DR InterPro: IPR009074; Apolipo_A_E_C3.
KW Cholesterol metabolism; Direct protein sequencing; HDL;
KW Lipid transport; Plasma.
FT NON TER 20
FT 20
SQ SEQUENCE 20 AA; 2387 MW; 9C970997C7FC976A CRC64;

Query Match 11.4%; Score 101; DB 1; Length 20;
Best Local Similarity 85.0%; Pred. No. 5.3;
Matches 17; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DEPPQSPWDRVKDLATVYVD 20
Db 1 DEPPQTPWDRVKDLVTVYVE 20
|||||:|||||:|||||:
|||||:|||||:|||||:

RESULT 14
Q9QV04 PRELIMINARY; PRT; 30 AA.
ID Q9QV04 AC Q9QV04;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Apolipoprotein E (Fragment).
OS Rattus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10118;
[1] RN
RP SEQUENCE
RX MEDLINE=95355312; PubMed=7629028;
RA Motojima K., Goto S.;
RL J. Biochem. 117:597-602(1995).
DR GO: GO:0005576; C:extracellular; IEA.
DR GO: GO:0008289; F:lipid binding; IEA.
DR GO: GO:0008689; P:lipid transport; IEA.
DR GO: GO:0042157; P:lipoprotein metabolism; IEA.
DR InterPro: IPR000074; Apolipoprotein.
DR Pfam: PF01442; Apolipoprotein; 1.
SQ SEQUENCE 30 AA; 3355 MW; 4A0103DCB6242BF3 CRC64;

Query Match 11.4%; Score 101; DB 2; Length 30;
Best Local Similarity 67.9%; Pred. No. 8.1;
Matches 19; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

Qy 4 PQSPWDRVKDLATVYVDLKGSDRDYVS 31
|||||:|||||:|||||:
Db 3 PQSQWDXVDFAIVYVDVXDGSXDYXS 30
|||||:|||||:|||||:

RESULT 15
APE_SAISC STANDARD; PRT; 107 AA.
ID APE_SAISC AC Q28995;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Apolipoprotein E (Apo-E) (Fragment).
GN Name=APOE;
OS Saimiri sciureus (Common squirrel monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Cebinae; Saimiri.
OX NCBI_TaxID=9521;
[1] RN
RP SEQUENCE FROM N.A.
RX MEDLINE=96225955; PubMed=8635577;
RA Morelli L., Wei L., Amorim A., McDermid J., Abec C.R., Frangione B.,
RA Walker L.C., Levy E.;

```

```

RT "Cerebrovascular amyloidosis in squirrel monkeys and rhesus monkeys:
RT apolipoprotein E genotype.";
RL FEBS Lett. 379:132-134(1996).
CC -!- FUNCTION: Mediates the binding, internalization, and catabolism of
CC lipoprotein particles. It can serve as a ligand for the LDL (apo
CC B/E) receptor and for the specific apo-E receptor (chylomicron
CC remnant) of hepatic tissues.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Secreted in plasma.
CC -!- SIMILARITY: Belongs to the apolipoprotein A1/A4/E family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U52029; AAC50442.1; -.
DR HSP; P02649; 1NFN.
DR InterPro; IPR000074; Apolipoprotein.
DR Pfam; PF01442; Apolipoprotein; 1.
KW Chylomicron; Glycoprotein; HDL; Heparin-binding; Lipid transport;
KW Plasma; Repeat; VLDL.
FT NON_TER 1
FT DOMAIN 89 99 LDL receptor binding (Potential).
FT DOMAIN 93 96 Heparin-binding (By similarity).
FT DOMAIN 11 >107 8 X 22 AA approximate tandem repeats.
FT REPEAT 11 32 1.
FT REPEAT 33 54 2.
FT REPEAT 55 76 3.
FT REPEAT 77 98 4.
FT REPEAT 99 >107 5.
FT NON_TER 107 107
SQ SEQUENCE 107 AA; 12328 MW; FF88CED47BD18F7C CRC64;
Query Match 11.2%; Score 99; DB 1; Length 107;
Best Local Similarity 26.2%; Pred. No. 41;
Matches 27; Conservative 19; Mismatches 35; Indels 22; Gaps 1;
QY 58 SKLRQLGQVTOEFWDNLEKETEGLRQEMSKDLVEVKAKVQPYLDDFQKKQOEEMELYRQ 117
Db 25 SELEQLSPVAETETARLSKELQAAQARLGADMDVRSRLAQY----- 67
QY 118 KVEPLRAELQEGARQKLHELOEKLSPGLGEMMDRARAHDALR 160
Db 68 -----RSEVQAMLGQSTDELRLARLASHLRKLRLLRDVDDLIQ 105

```

Search completed: December 21, 2004, 13:51:30
Job time : 190 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2004, 10:45:04 ; Search time 66 Seconds
(without alignments)
924.000 Million cell updates/sec

Title: US-09-803-918A-2_COPY_25_194

Perfect score: 885
Sequence: 1 DEPPQSPWRVKDLATYYVD.....RAAHVADLRTHLAPYSDEL 170

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 segs, 358729299 residues

Total number of hits satisfying chosen parameters: 1323736

Minimum DB seq length: 0
Maximum DB seq length: 170

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	885	100.0	170	6	ADA61215	Ada61215 Human 18K
2	727	82.1	151	4	AAO02278	Aao02278 Human pol
3	647.5	73.2	166	4	AAU28372	Aau28372 Novel hum
4	612	69.2	160	6	ABP75964	Abp75964 Human GEN
5	529.5	59.8	168	5	AAE24649	Aae24649 Human MSP
6	501.5	56.7	154	4	AAO12095	Aao12095 Human pol
7	479.5	54.2	119	4	AAU30468	Aau30468 Novel hum
8	476.5	53.8	120	4	AAU30469	Aau30469 Novel hum
9	476.5	53.8	120	4	AAU30267	Aau30267 Novel hum
10	442.5	50.0	168	5	AAE24648	Aae24648 Human MSP
11	437	49.4	158	7	ADJ70395	Adj70395 Human hea
12	218	24.6	42	2	AAR20164	Aar20164 Apo AI po
13	197	22.3	65	6	ABP76131	Abp76131 Human GEN
14	163	18.4	32	2	AAR20165	Aar20165 Apo AI po
15	162	18.3	64	5	ABP31744	Abp31744 Human apo
16	138	15.6	26	6	AAG62609	Aag62609 Apolipop
17	137	15.5	26	7	ADE76474	Ade76474 ApoAI G*
18	123	13.9	85	8	ADP87441	Adp87441 Monkey ap
19	119	13.4	85	8	ADP87442	Adp87442 Human apo
20	119	13.4	85	8	ADP87440	Adp87440 Baboon ap
21	118	13.3	107	1	AAP92072	Aap92072 Apolipop
22	118	13.3	107	8	ADP87432	Adp87432 Rat apoli
23	116	13.1	21	1	AAP90956	Aap90956 Apo AI ep
24	115	13.0	107	8	ADP87431	Adp87431 Rat apoli
25	111	12.5	22	4	AAG62608	Aag62608 Apolipop

ALIGNMENTS

RESULT 1									
ADA61215									
ID	ADA61215	standard;	protein;	170	AA.				
XX	ADA61215;								
XX									
DT	20-NOV-2003	(first entry)							
XX									
DE	Human 18K apolipoprotein A-1 (apo-A-1) fragment.								
XX									
KW	Human; apolipoprotein A-I; apo-A-I; T-cell activation inhibitor-like;								
KW	AFPI; T-cell-mediated activation; monocyte; TNF-alpha; interleukin-beta; IL-1beta;								
KW	tumour necrosis factor alpha; TNF-alpha; gene therapy;								
KW	IL-1-mediated disease; acute pancreatitis; Alzheimer's disease; asthma;								
KW	cancer; diabetes; glomerulonephritis; multiple sclerosis; osteoporosis;								
KW	pain; Parkinson's disease; psoriasis; uveitis; TNF-mediated disease;								
KW	depression; pancreatitis; periodontal disease; pulmonary fibrosis;								
KW	reperfusion injury; rheumatic disease; septic shock;								
KW	systemic lupus erythematosus; thyroiditis; nontropic; neuroprotective;								
KW	cytostatic; nephrotropic; analgesic; ophthalmological; antiinflammatory;								
KW	vasotropic; antibacterial; immunosuppressive; dermatological.								
XX									
OS	Homo sapiens.								
XX									
PN	US2002064820-A1.								
XX									
PD	30-MAY-2002.								
XX									
PF	13-MAR-2001; 2001US-00803918.								
XX									
PR	13-MAR-2000; 2000US-0189008P.								
XX									
PR	31-MAR-2000; 2000US-0193551P.								
XX									
PA	(DAYE/) DAYER J.								
PA	(BURGH/) BURGER D.								
PA	(KOHNO/) KOHNO T.								
XX	(EDWA/) EDWARDS C K.								
PI	Dayer J, Burger D, Kohno T, Edwards CK;								
DR	WPI; 2003-615701/58.								
DR	N-PSDB; ADA61216.								
XX									
PT	New apolipoprotein A-I polypeptide or its related polypeptide termed apo-								
PT	A-I fragment T-cell activation inhibitor-like polypeptide, useful for								
XX	treating interleukin-1 or tumor necrosis factor-alpha mediated disease.								
PS	Claim 15; Fig 1B; 56pp; English.								

XX The invention discloses a apolipoprotein A-I (apo-A-I) polypeptide or its
 CC related polypeptide, apo-A-I fragment T-cell activation inhibitor-like
 CC (AFTI) polypeptide, or a fragment or an allelic or splice variant. Apo-A-
 CC I and its derivatives are used in regulating T-cell-mediated activation
 CC of monocytes. Also disclosed is an antibody produced by immunising an
 CC animal with apo-A-I and a selective binding agent, or its fragment. Apo-A
 CC -I is useful for reducing inflammation, interleukin (IL)-1beta secretion
 CC or tumour necrosis factor (TNF)-alpha secretion in a subject, or for
 CC treating IL-1 or TNF-alpha mediated disease in a subject. The
 CC compositions of the invention are useful for treating (e.g. gene
 CC therapy), preventing or ameliorating a medical condition involving
 CC monocyte activation in a patient, for determining whether a compound
 CC inhibits activity or production of apo-A-I, for detecting or quantifying
 CC the amount of apo-A-I in a sample, for regulating T-cell-mediated
 CC activation of monocytes, to identify a molecule that binds to apo-A-I,
 CC for identifying AFTI receptors, for cloning AFTI receptors, and as an
 CC immunogen to raise antibodies, for treating acute and chronic IL-1-
 CC mediated diseases such as acute pancreatitis, Alzheimer's disease,
 CC asthma, cancer, diabetes, glomerulonephritis, multiple sclerosis,
 CC osteoporosis, pain, Parkinson's disease, psoriasis, uveitis, and acute
 CC and chronic TNF-mediated diseases such as depression, pancreatitis,
 CC periodontal diseases, pulmonary fibrosis, reperfusion injury, rheumatic
 CC diseases, septic shock, systemic lupus erythematosus and thyroiditis.
 CC They are also useful for regulating expression and modulating levels of
 CC apo-A-I, as hybridisation probes in diagnostic assays, for diagnostic
 CC applications, to test for the presence of apo-A-I in cells, for in vivo
 CC imaging and in therapeutics. The sequence presented is the human 18k N-
 CC terminal apo-A-I protein fragment.

XX Sequence 170 AA;

Query Match 100.0%; Score 885; DB 6; Length 170;
 Best Local Similarity 100.0%; Pred. No. 8.8e-66;
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DBPPQSPWRVKDQATVYDVVLKDSGRDYVSQFEGSALGKQLNLKLLDNWDSVTSTFSKL 60
 DB 1 DBPPQSPWRVKDQATVYDVVLKDSGRDYVSQFEGSALGKQLNLKLLDNWDSVTSTFSKL 60

QY 61 RQQLGPTQTFWMDNLEKETEGRLQEMSKDLEEVKAKVQPYLDDFQKKWQEEMLYRQKVE 120
 DB 61 RQQLGPTQTFWMDNLEKETEGRLQEMSKDLEEVKAKVQPYLDDFQKKWQEEMLYRQKVE 120

QY 121 PIRAELOEGARQKLHELOEKLSPLGEMDRARAHVDALRTHLAPYSDEL 170
 DB 121 PIRAELOEGARQKLHELOEKLSPLGEMDRARAHVDALRTHLAPYSDEL 170

RESULT 2
 AAO02278
 ID AAO02278 standard; protein; 151 AA.

XX AAO02278;

XX 06-NOV-2001 (first entry)

DE Human polypeptide SEQ ID NO 16170.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorders; arthritis; inflammation.

OS Homo sapiens.

XX WO200164835-A2.

XX 07-SEP-2001.

XX 26-FEB-2001; 2001WO-US004927.

XX 28-FEB-2000; 2000US-00515126.

PR 18-MAY-2000; 2000US-00577409.

XX (HYSE-) HYSEQ INC.

PA Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-514838/56.

DR N-PSDB; AAI82209.

XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing
 PT and treating e.g. leukemia, inflammation and immune disorders.

XX Claim 20; SEQ ID NO 16170; 1399pp + Sequence Listing; English.

XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
 CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation. Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 151 AA;

Query Match 82.1%; Score 727; DB 4; Length 151;
 Best Local Similarity 94.0%; Pred. No. 1.1e-52;
 Matches 140; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 22 LKDSGRDYVSQFEGSALGKQLNLKLLDNWDSVTSTFSKLRLQGLGPTQTFWMDNLEKETEG 81
 DB 1 LKDSGRDYVSQFEGSALGKQLNLKLLDNWDSVTSTFSKLRLQGLGPTQTFWMDNLEKETEG 60

QY 82 LRQEMSKDLEEVKAKVQPYLDDFQKKWQEEMLYRQKVEPIRAELOEGARQKLHELOEKL 141
 DB 61 LRQEMSKDLEEVKAKVQPYLDDFQKKWQEEMLYRQKVEPIRAELOEGARQKLHELOEKL 120

QY 142 SPLGEMDRARAHVDALRTHLAPYSDEL 170
 DB 121 SPLGEMDRARAHVDALRTHLAPYSDEL 149

RESULT 3
 AAU28372
 ID AAU28372 standard; protein; 166 AA.

XX AAU28372;

XX 18-DEC-2001 (first entry)

XX Novel human secretory protein, Seq ID No 729.

XX Human; secreted protein; arthritis; Crohn's disease; sepsis; shock;
 KW ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy;
 KW transgenic animal; Alzheimer's disease; Parkinson's disease; burn;
 KW amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;
 KW ulcer; osteoporosis; bone degenerative disorder; periodontal disease;
 KW gut protection; lung; liver fibrosis; immune deficiency; infection;
 KW severe combined immunodeficiency; SCID; autoimmune disorder; allergy;
 KW multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;
 KW fertility; analgesic; pain; antigen.

XX Homo sapiens.

XX WO200166689-A2.

XX 13-SEP-2001.

PF 05-MAR-2001; 2001WO-US004942.
XX
XX 07-MAR-2000; 2000US-00519705.
PR 19-MAY-2000; 2000US-00574454.
PR 17-JUN-2000; 2000US-00596193.
PR 14-JUL-2000; 2000US-00618847.
PR 19-SEP-2000; 2000US-00665363.
PR 20-OCT-2000; 2000US-00693267.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Asundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P;
PI Zhao QA, Yang Y, Drmanac RT, Zhang J, Chen R, Xue AJ, Wang J;
XX
XX WPI; 2001-589934/66.
DR N-PSDB; AAS45272.
XX
XX Novel polypeptides and nucleic acids obtained from cDNA libraries
PT prepared from various human tissues, for diagnosis and treatment of
PT cancer, neurological, inflammatory, and autoimmune disorders.
XX
XX Example 2; SEQ ID NO 729; 107pp; English.
XX
XX The invention relates to novel isolated human secreted polypeptides (I)
CC and polynucleotides (II). (I) and (II) are useful for treating
CC inflammatory conditions such as arthritis, nephritis, Crohn's disease,
CC ischaemia-reperfusion injury, shock, sepsis, immune responses, and is
CC involved in increasing haematopoiesis, stem cell survival, bone growth
CC and remodeling. (I), (II) and modulators of (II) are useful for
CC prophylaxis or treatment of one or more cancers. (II) is also useful for
CC creating transgenic animals useful for studying the in vivo activities of
CC the polypeptide as well as for studying modulators of the polypeptides.
CC (I) induces the proliferation of neural cells and regeneration of nerve
CC and brain tissue and is useful for the treatment of central and
CC peripheral nervous system diseases and neuropathies, such as Alzheimer's,
CC Parkinson's disease, Huntington's disease, and amyotrophic lateral
CC sclerosis. In addition, (I) is involved in chemotactic or chemokinetic
CC activity, regulation of haematopoiesis and is useful for treating myeloid
CC or lymphoid cell disorders, platelet disorders such as thrombocytopenia
CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve
CC tissue growth, and in tissue repair, healing of burns, incisions, ulcers,
CC for treating osteoporosis, osteoarthritis, bone degenerative disorders,
CC or periodontal disease. Furthermore, (I) is also useful for gut
CC protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues, various immune deficiencies and
CC disorders including severe combined immunodeficiency (SCID), bacterial or
CC fungal infections, autoimmune disorders e.g. multiple sclerosis,
CC rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic
CC reactions and conditions, such as asthma or other respiratory problems.
CC In addition, (I) affects biorhythms or circadian cycles of rhythms,
CC fertility, metabolism, catabolism, anabolism, storage or elimination of
CC dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides
CC analgesic effects or other pain reducing effects, immunoglobulin like
CC activity and can act as an antigen in a vaccine composition to raise an
CC immune response. AAU28020-AAU28395 represent novel human secreted protein
CC amino acid sequences of the invention
XX
XX Sequence 166 AA;
Query Match 73.2%; Score 647.5; DB 4; Length 166;
Best Local Similarity 82.7%; Pred. No. 5.1e-46;
Matches 129; Conservative 1; Mismatches 3; Indels 23; Gaps 1;
QY 1 DEPPQSPWDRVKDLATVYVDVLKSGRDYVSQFEGSALGKQLNLKLLDNWDSVTSTFSKL 60
DB 34 DEPPQSPWDRVKDLATVYVDVLKSGK-----DSVTSTFSKL 70
QY 61 REQLGPTVQEFWDNLEKETEGRLQEMSKDLEEVKAKVQPYLDDFQKKWQEMELYRQKYE 120
DB 71 REQLGPTVQEFWDNLEKETEGRLQEMSKDLEEVKAKVQPYLDDFQKKWQEMELYRQKYE 130
QY 121 PLRAELQEGARQKLHELOEKLSPLGEEADRRRAHV 156
|||||

DB 131 PLRAELQEGARQKLHELOEKLSPLGEEADRRRAHV 166
RESULT 4
ABP75964
ID ABP75964 standard; protein; 160 AA.
XX
XX AC ABP75964;
XX
XX 21-PBB-2003 (first entry)
XX
XX Human GENSET protein SEQ ID 171.
XX
XX Cytostatic; antiinflammatory; nootropic; neuroprotective; cardiant;
KW gastrointestinal; gene therapy; GENSET; heavy metal toxicity; cancer;
KW inflammatory disease; immune disorder; neuromuscular; toxicity;
KW central nervous system; cardiovascular; gastrointestinal.
XX
XX OS Homo sapiens.
XX
XX WO200283898-A1.
XX
XX 24-OCT-2002.
XX
XX 18-APR-2001; 2001WO-IB0000914.
XX
XX 18-APR-2001; 2001WO-IB0000914.
XX
XX (GEST) GENSET.
XX
XX Bejanin S, Tanaka H, Dumas Milne Edwards J, Jobert S, Giordano J;
PI WPI; 2003-075548/07.
XX
XX New GENSET polynucleotides and polypeptides, useful for treating heavy
PT metal toxicity, cancer, inflammatory diseases, immune disorders, and the
PT neuromuscular, CNS, cardiovascular or gastrointestinal effects of the
PT toxicity.
XX
XX Claim 14; Page 402; 735pp; English.
XX
XX The present invention relates to novel GENSET polynucleotides (ABZ36404-
CC ABZ36911) encoding polypeptides (ABP75963-ABP76368). The polynucleotides
CC and polypeptides are useful in screening and diagnostic assays for
CC abnormal GENSET expression and/or biological activity. They are also
CC useful for screening of compounds for treating or preventing GENSET-
CC related disorders, such as heavy metal toxicity, cancer, inflammatory
CC diseases, immune disorders, and the neuromuscular, central nervous system
CC (CNS), cardiovascular or gastrointestinal effects of the toxicity
XX
XX Sequence 160 AA;
Query Match 69.2%; Score 612; DB 6; Length 160;
Best Local Similarity 98.3%; Pred. No. 4.4e-43;
Matches 116; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 DEPPQSPWDRVKDLATVYVDVLKSGRDYVSQFEGSALGKQLNLKLLDNWDSVTSTFSKL 60
DB 25 DEPPQSPWDRVKDLATVYVDVLKSGRDYVSQFEGSALGKQLNLKLLDNWDSVTSTFSNL 84
QY 61 REQLGPTVQEFWDNLEKETEGRLQEMSKDLEEVKAKVQPYLDDFQKKWQEMELYRQK 118
DB 85 REQLGPTVQEFWDNLEKETEGRLQEMSKDLEEVKAKVQPYLDDFQKKWQEMELYRQK 142
RESULT 5
AAE24649
ID AAE24649 standard; protein; 168 AA.
XX
XX AC AAE24649;
XX
XX 22-OCT-2002 (first entry)
XX

DE Human MSP1D5-6 (helices 5 and 6 deleted) protein.

XX Human; membrane scaffold protein; MSP; phospholipid; nanoscale particle;

KW hydrogel; gold biosensor surface; reproducible assay; crystallisation;

KW biotechnology; pharmaceutical industry; structure determination;

KW bioseparation; drug discovery.

XX Homo sapiens.

OS Synthetic.

XX WO200240501-A2.

XX 23-MAY-2002.

XX 20-NOV-2001; 2001WO-US043451.

XX 20-NOV-2000; 2000US-0252233P.

XX (UNII) UNIV ILLINOIS FOUND.

XX Sligar SG, Bayburt TH;

XX WPI; 2002-500201/53.

DR N-PSDB; AAD39711.

XX New membrane scaffold protein for forming nanoscale particles, useful in

PT biological research, self assemblies without) phospholipids into a

PT nanoscale particle in an aqueous environment.

XX Claim 8; Page 39; 112pp; English.

XX The invention relates to a membrane scaffold protein (MSP), that self

CC assembles in the absence of phospholipid, with a phospholipid or a

CC mixture of phospholipids, into a nanoscale particle of 5-500 nm in

CC diameter, in an aqueous environment, where the membrane scaffold protein

CC is amphipathic and forms at least one alpha helix. The nanoscale particle

CC comprising MSP is useful for identifying a competitor of binding of a

CC ligand to a receptor protein which is incorporated within the nanoscale

CC particle together with MSP, by contacting nanoscale particle and a

CC receptor protein with a detectable ligand to produce nanoscale particle-

CC bound detectable ligand, contacting the bound ligand with a test

CC compound, and measuring detectable ligand released from the particle,

CC such that a competitor of ligand binding is identified when contacting

CC the bound ligand results in release of the detectable ligand. The

CC invention is useful for the incorporation of additional hydrophobic or

CC partially hydrophobic protein molecules. The nanoscale particle can form

CC tags for purification and physical manipulation of disks such as in

CC hydrogels on or a gold biosensor surface, and they can serve as entities

CC for rapid and reproducible assay and crystallisation. The nanoscale

CC particle and MSP are useful in biotechnology, pharmaceutical industries

CC and in research areas. The nanoscale particle is useful for structure/

CC function correlation, structure determination, bioseparation and drug

CC discovery. The present sequence is MSP1D5-6 protein

XX Sequence 168 AA;

SQ Query Match 59.8%; Score 529.5; DB 5; Length 168;

Best Local Similarity 74.5%; Pred. No. 3.4e-36;

Matches 108; Conservative 5; Mismatches 25; Indels 7; Gaps 1;

QY 26 GRDYSQFSGSALGKQLNLKLDNWDVSTSTSKLREQLGPTVQFWDNLKETGLRQE 85

2 GHHHHHIG- - - - - RUKLDNWDVSTSTSKLREQLGPTVQFWDNLKETGLRQE 54

86 MSKDLSEVAKVQPYLDDDFQKKWQEMELYRQKVEPLRAELQEGARQKLHEQKLSPLG 145

55 MSKDLSEVAKVQPYLDDDFQKKWQEMELYRQKVEPLRAELQEGARQKLHEQKLSARL 114

146 EWRDRARAHVADALRTHLAPYSDEL 170

DB 115 AEYHAKATEHLSTLSEKAKPALEDL 139

RESULT 6

AAO12095

ID AAO12095 standard; protein; 154 AA.

XX AAO12095;

XX 06-NOV-2001 (first entry)

XX Human polypeptide SEQ ID NO 25987.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;

KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;

KW tissue growth factor; immunomodulatory; cancer; leukaemia;

KW nervous system disorders; arthritis; inflammation.

XX Homo sapiens.

XX WO200164835-A2.

XX 07-SEP-2001.

XX 26-FEB-2001; 2001WO-US004927.

XX 28-FEB-2000; 2000US-00515126.

PR 18-MAY-2000; 2000US-00577409.

XX (HYSE-) HYSEQ INC.

PA Tang YT, Liu C, Drmanac RT;

PI WPI; 2001-514838/56.

DR N-PSDB; AAI92026.

XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing

PT and treating e.g. leukemia, inflammation and immune disorders.

XX Claim 20; SEQ ID NO 25987; 1399pp + Sequence Listing; English.

XX The invention relates to human polynucleotides (AAI79941-AAI93841) and

CC the encoded proteins (AAO0010-AAO13910) that exhibit activity relating to

CC cytokine, cell proliferation or cell differentiation or which may induce

CC production of other cytokines in other cell populations. The

CC polynucleotides and polypeptides are useful in gene therapy, vaccines or

CC peptide therapy. The polypeptides have various cytokine-like activities,

CC e.g. stem cell growth factor activity, haematopoiesis regulating

CC activity, tissue growth factor activity, immunomodulatory activity and

CC activin/inhibin activity and may be useful in the diagnosis and/or

CC treatment of cancer, leukaemia, nervous system disorders, arthritis and

CC inflammation. Note: The sequence data for this patent did not form part

CC of the printed specification, but was obtained in electronic format

CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 154 AA;

SQ Query Match 56.7%; Score 501.5; DB 4; Length 154;

Best Local Similarity 79.5%; Pred. No. 6.6e-34;

Matches 97; Conservative 7; Mismatches 17; Indels 1; Gaps 1;

QY 1 DEPPSQPDRVVKDLATVVDVLKDSGRDYVSQFEGSALGKQLNLKLDNWDVSTSTFSKL 60

33 DEPPSQPDRVMDLATAVVDGLKDSGRDYASQFEGYALGKQLNLMLLDNWDVSTVTFSKL 92

61 REQLGPTVQFWDNLKETGLRQEMSKDLSEVAKVQPYLDDDFQKKWQEMELYRQKVE 120

93 REQLGPTVQFWDNLKETGLRQEMSKDLSEVAKVQPYLDDDFQKKWQEMELYRQKVE 151

121 PL 122

152 PV 153

RESULT 7

AAU30468

ID AAU30468 standard; protein; 119 AA.
XX
AC AAU30468;
XX
DT 18-DEC-2001 (first entry)
XX
DE Novel human secreted protein #959.
XX
KW Human; vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX
OS Homo sapiens.
XX
PN WO200179449-A2.
XX
PD 25-OCT-2001.
XX
PF 16-APR-2001; 2001WO-US008656.
XX
PR 18-APR-2000; 2000US-00552929.
PR 26-JAN-2001; 2001US-00770160.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX WPI; 2001-611725/70.
XX
PT Nucleic acids encoding a range of human polypeptides, useful in genetic
PT vaccination, testing and therapy.
XX
PS Claim 20; Page 297; 765pp; English.
XX
CC The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising the
CC nucleic acids encoding the polypeptides and cells genetically engineered
CC to express them are also useful for producing the proteins. The proteins
CC are useful in genetic vaccination, testing and therapy, and can be used
CC as nutritional supplements. They may be used to increase stem cell
CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
CC and/or nerve tissue growth or regeneration; immune suppression and/or
CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
CC AAU29510-AAU33304 represent the amino acid sequences of novel human
CC secreted proteins of the invention
XX
SQ Sequence 119 AA;
Query Match 54.2%; Score 479.5; DB 4; Length 119;
Best Local Similarity 89.0%; Pred. No. 3.3e-32;
Matches 97; Conservative 1; Mismatches 6; Indels 5; Gaps 1;
QY 43 NLKLLDNWDSVTSFSLKREQLGPVTQEFWDNLEKETEGLRQMSKDLSEVKAKVQPYLD 102
DB 2 NLKLLDNWDSVTSFSLKREQLGPVTQEFWDNLEKETEGLRQMSKDLSEVKAKVQPYLD 61
QY 103 DFQKKQWEMELYRQKVEPLRAELQEGARQKLHEQK-----LSPLGE 146
DB 62 DFQKKQWEMELYRQKVEPLRAELQEGARQKLPLVLESFKVSFLSALE 110
RESULT 8
AAU30469
ID AAU30469 standard; protein; 120 AA.
XX
AC AAU30469;
XX

DT 18-DEC-2001 (first entry)
XX
DE Novel human secreted protein #960.
XX
KW Human; vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX
OS Homo sapiens.
XX
PN WO200179449-A2.
XX
PD 25-OCT-2001.
XX
PF 16-APR-2001; 2001WO-US008656.
XX
PR 18-APR-2000; 2000US-00552929.
PR 26-JAN-2001; 2001US-00770160.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX WPI; 2001-611725/70.
XX
PT Nucleic acids encoding a range of human polypeptides, useful in genetic
PT vaccination, testing and therapy.
XX
PS Claim 20; Page 297; 765pp; English.
XX
CC The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising the
CC nucleic acids encoding the polypeptides and cells genetically engineered
CC to express them are also useful for producing the proteins. The proteins
CC are useful in genetic vaccination, testing and therapy, and can be used
CC as nutritional supplements. They may be used to increase stem cell
CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
CC and/or nerve tissue growth or regeneration; immune suppression and/or
CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
CC AAU29510-AAU33304 represent the amino acid sequences of novel human
CC secreted proteins of the invention
XX
SQ Sequence 120 AA;
Query Match 53.8%; Score 476.5; DB 4; Length 120;
Best Local Similarity 79.7%; Pred. No. 5.9e-32;
Matches 94; Conservative 1; Mismatches 0; Indels 23; Gaps 1;
QY 1 DEPPQSPWDRVKDLATVYVDVLKSGRDYVSQFEGSALGKQLNLKLLDNWDSVTSFSLK 60
DB 25 DEPPQSPWDRVKDLATVYVDVLKDSGK-----DSVTSTFSLK 61
QY 61 REQLGPVTQEFWDNLEKETEGLRQMSKDLSEVKAKVQPYLDDFKKKQWEMELYRQK 118
DB 62 REQLGPVTQEFWDNLEKETEGLRQMSKDLSEVKAKVQPYLDDFKKKQWEMELYRQK 119
RESULT 9
AAU30267
ID AAU30267 standard; protein; 120 AA.
XX
AC AAU30267;
XX
DT 18-DEC-2001 (first entry)
XX
DE Novel human secreted protein #758.
XX

KW Human; vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX Homo sapiens.
OS WO200179449-A2.
PN 25-OCT-2001.
XX 16-APR-2001; 2001WO-US008656.
XX 18-APR-2000; 2000US-00552929.
PR 26-JAN-2001; 2001US-00770160.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Drmanac RT;
XX WPI; 2001-611725/70.
DR Nucleic acids encoding a range of human polypeptides, useful in genetic
PT vaccination, testing and therapy.
XX Claim 20; Page 270; 765pp; English.
XX The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising the
CC nucleic acids encoding the polypeptides and cells genetically engineered
CC to express them are also useful for producing the proteins. The proteins
CC are useful in genetic vaccination, testing and therapy, and can be used
CC as nutritional supplements. They may be used to increase stem cell
CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
CC and/or nerve tissue growth or regeneration; immune suppression and/or
CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
CC AAU29510-AAU33304 represent the amino acid sequences of novel human
CC secreted proteins of the invention
XX Sequence 120 AA;
SQ
Query Match 53.8%; Score 476.5; DB 4; Length 120;
Best Local Similarity 79.7%; Pred. No. 5.9e-32;
Matches 94; Conservative 1; Mismatches 0; Indels 23; Gaps 1;
QY 1 DEPPQSPWDRVKDLATVYVDVLKDSGRDYVSQFEGSALGKQLNLKLLDNWDSVTSTFSKL 60
DB 25 DEPPQSPWDRVKDLATVYVDVLKDSGK-----DSVTSTFSKL 61
QY 61 RQLGPGVTOEFWDNLEKETEGRLQEMSKDLEEVKAKVQPYLDDFOKKQOEEMELYRQK 118
DB 62 RQLGPGVTOEFWDNLEKETEGRLQEMSKDLEEVKAKVQPYLDDFOKKQOEEMELYRQK 119
RESULT 10
AAE24648
ID AAE24648 standard; protein; 168 AA.
XX AAE24648;
XX 22-OCT-2002 (first entry)
XX Human MSP1d4-5 (helices 4 and 5 deleted) protein.
XX Human; membrane scaffold protein; MSP; phospholipid; nanoscale particle;
KW hydrogel; gold biosensor surface; reproducible assay; crystallisation;
KW biotechnology; pharmaceutical industry; structure determination;
KW bioprocessing; drug discovery.

XX Homo sapiens.
OS Synthetic.
XX WO200240501-A2.
PN 23-MAY-2002.
XX 20-NOV-2001; 2001WO-US043451.
XX 20-NOV-2000; 2000US-0252233P.
XX (UNITI) UNIV ILLINOIS FOUND.
XX Sligar SG, Bayburt TH;
PI WPI; 2002-500201/53.
DR N-PSDB; AAD39706.
XX New membrane scaffold protein for forming nanoscale particles, useful in
PT biological research, self assemblies with(out) phospholipids into a
PT nanoscale particle in an aqueous environment.
XX Claim 8; Page 38; 112pp; English.
XX The invention relates to a membrane scaffold protein (MSP), that self
CC assembles in the absence of phospholipid, with a phospholipid or a
CC mixture of phospholipids, into a nanoscale particle of 5-500 nm in
CC diameter, in an aqueous environment, where the membrane scaffold protein
CC is amphipathic and forms at least one alpha helix. The nanoscale particle
CC comprising MSP is useful for identifying a competitor of binding of a
CC ligand to a receptor protein which is incorporated within the nanoscale
CC particle together with MSP, by contacting nanoscale particle and a
CC receptor protein with a detectable ligand to produce nanoscale particle-
CC bound detectable ligand, contacting the bound ligand with a test
CC compound, and measuring detectable ligand released from the particle,
CC such that a competitor of ligand binding is identified when contacting
CC the bound ligand results in release of the detectable ligand. The
CC invention is useful for the incorporation of additional hydrophobic or
CC partially hydrophobic protein molecules. The nanoscale particle can form
CC tags for purification and physical manipulation of disks such as in
CC hydrogels on or a gold biosensor surface, and they can serve as entities
CC for rapid and reproducible assay and crystallisation. The nanoscale
CC particle and MSP are useful in biotechnology, pharmaceutical industries
CC and in research areas. The nanoscale particle is useful for structure/
CC function correlation, structure determination, bioprocessing and drug
CC discovery. The present sequence is MSP1D4-5
XX Sequence 168 AA;
SQ
Query Match 50.0%; Score 442.5; DB 5; Length 168;
Best Local Similarity 63.4%; Pred. No. 6e-29;
Matches 92; Conservative 9; Mismatches 37; Indels 7; Gaps 1;
QY 26 GRDYVSQFEGSALGKQLNLKLLDNWDSVTSTFSKLREQLGPGVTOEFWDNLEKETEGRLQ 85
DB 2 GHHHHHIEG-----RLKLLDNWDSVTSTFSKLREQLGPGVTOEFWDNLEKETEGRLQ 54
QY 86 MSKDLLEEVKAKVQPYLDDFOKKQOEEMELYRQKVEPLRAELQEGARQKHLQEKLSPLG 145
DB 55 MSKDLLEEVKAKVQPYLDDFOKKQOEEMELYRQKVEPLRAELQEGARQKHLQEKLSPLG 144
QY 146 EEMDRARAHDALATHLAPYSDEL 170
DB 115 AEYHAKATEHLSTLSEKAKPALEDL 139
RESULT 11
ADJ70395
ID ADJ70395 standard; protein; 158 AA.
XX ADJ70395;
XX

DT XX 06-MAY-2004 (first entry)
DE XX Human heart mitochondrial protein as a therapeutic target SeqID2201.
XX
XX mitochondrial; human; screening assay; diabetes mellitus;
KW Huntington's disease; osteoarthritis;
KW Leber's hereditary optic neuropathy; LHON;
KW mitochondrial encephalopathy lactic acidosis and stroke; MELAS;
KW myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;
KW neuroprotective; neurotropic; antidiabetic; anticonvulsant; antiarthritic;
KW osteopathic; ophthalmological; cytostatic.
XX
OS Homo sapiens.
XX
XX WO2003087768-A2.
PN
XX
XX 23-OCT-2003.
PD
XX
XX 04-APR-2003; 2003WO-US010870.
PF
XX 12-APR-2002; 2002US-0372843P.
PR
XX 17-JUN-2002; 2002US-0389987P.
PR
XX 20-SEP-2002; 2002US-0412418P.
PR
XX
XX (MITO-) MITOKOR.
PA
PA (BUCK-) BUCK INST AGE RES.
XX
XX Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM;
PI Warnock DE;
PI
XX
XX WPI; 2003-845369/78.
DR
XX
XX Identifying a mitochondrial target for drug screening assays and for
PT treating diseases associated with altered mitochondrial function,
PT comprise detecting a modified polypeptide in a sample and correlating
PT with the disease.
PT
XX
XX Claim 1; SEQ ID NO 2201; 180pp; English.
PS
XX
XX This invention relates to novel mitochondrial targets that can be used
CC for therapeutic intervention in treating a disease associated with
CC altered mitochondrial function. Specifically, it refers to a method for
CC identifying proteins of the human heart mitochondrial proteome that are
CC useful for drug screening assays, as well as therapeutic targets. The
CC present invention describes a method for identifying such proteins that
CC can be used in the treatment of various diseases associated with altered
CC mitochondrial function including diabetes mellitus, Huntington's disease,
CC osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial
CC encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy
CC ragged red fibre syndrome (MERRF) or cancer. Accordingly, these
CC compositions have neuroprotective, neurotropic, antidiabetic,
CC anticonvulsant, antiarthritic, osteopathic, ophthalmological and
CC cytostatic activities. This polypeptide sequence is a human heart
CC mitochondrial protein of the invention.
XX
XX Sequence 158 AA;
SQ
Query Match 49.4%; Score 437; DB 7; Length 158;
Best Local Similarity 100.0%; Pred. No. 1.6e-28;
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 86 MSKDLSEVAKVQPYLDDFKKKQWEEMLYRQKVEPLRAEQKQLHELQEKLSPLG 145
DB 1 MSKDLSEVAKVQPYLDDFKKKQWEEMLYRQKVEPLRAEQKQLHELQEKLSPLG 60
QY 146 EEMRDRARAHVDALRTHLAPYSDEL 170
DB 61 EEMRDRARAHVDALRTHLAPYSDEL 85
RESULT 12
AAR20164
ID AAR20164 standard; peptide; 42 AA.

XX AAR20164;
AC
XX 25-MAR-2003 (revised)
DT 01-APR-1992 (first entry)
DT
XX Apo AI polypeptide (A).
DE
XX Monoclonal antibody; MAB AI-11; epitope; diagnosis.
KW Monoclonal antibody; MAB AI-11; epitope; diagnosis.
XX
XX Synthetic.
OS
XX
XX Key Location/Qualifiers
FH Peptide 1..28
FT Peptide 2..28
FT Peptide 7..28
FT Peptide 10..28
FT Peptide 11..42
FT Peptide 11..28
FT Peptide 13..28
XX
XX WO9118619-A.
PN
XX 12-DEC-1991.
PD
XX
XX 07-JUN-1990; 90US-00534761.
PF
XX
XX 07-JUN-1990; 90US-00534761.
PR
XX (SCRI) SCRIPPS CLINIC & RES.
PA
XX
XX Curtiss LK, Banka CL, Bonnet DJ, Smith RS;
PI
XX WPI; 1992-007201/01.
DR
XX New Apo AI polypeptide(s) and anti-Apo antibodies - used to detect Apo AI
PT in vascular fluid samples and increase LCAT-mediated cholesterol
PT esterification in humans.
PT
XX
XX Claim 2; Page 68; 87pp; English.
PS
XX The peptides represented in AAR20164-65 are capable of immunologically
CC mimicking an Apo AI epitope. They are useful in diagnosis and detection
CC of Apo AI in fluids and for preparing anti-Apo AI antibodies. It is
CC useful in therapeutic methods for increasing LCAT-mediated cholesterol
CC esterification in humans. The Apo AI polypeptide is selected from the
CC peptides indicated in the features and includes amino acids 13-28
CC defining a conserved native epitope on Apo AI capable of immunoreacting
CC with monoclonal antibody MAB AI-11. (Updated on 25-MAR-2003 to correct PA
CC field.)
XX
XX Sequence 42 AA;
SQ
Query Match 24.6%; Score 218; DB 2; Length 42;
Best Local Similarity 100.0%; Pred. No. 5.4e-11;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 84 QEMSKDLSEVAKVQPYLDDFKKKQWEEMLYRQKVEPLRAE 125
DB 1 QEMSKDLSEVAKVQPYLDDFKKKQWEEMLYRQKVEPLRAE 42
RESULT 13
ABP76131
ID ABP76131 standard; protein; 65 AA.
XX
XX AC ABP76131;
XX
XX 21-FEB-2003 (first entry)
DT
XX Human GENSET protein SEQ ID 457.
DE
XX Cytostatic; antiinflammatory; neurotropic; neuroprotective; cardiant;
KW

KW gastrointestinal; gene therapy; GENSET; heavy metal toxicity; cancer;
KW inflammatory disease; immune disorder; neuromuscular; toxicity;
KW central nervous system; cardiovascular; gastrointestinal.
XX Homo sapiens.
OS WO200283898-A1.
PN 24-OCT-2002.
PD
XX 18-APR-2001; 2001WO-IB000914.
XX 18-APR-2001; 2001WO-IB000914.
XX (GEST) GENSET.
PA
PI Bejanin S, Tanaka H, Dumas Milne Edwards J, Jobert S, Giordano J;
XX WPI; 2003-075548/07.
DR
XX New GENSET polynucleotides and polypeptides, useful for treating heavy
XX metal toxicity, cancer, inflammatory diseases, immune disorders, and the
XX neuromuscular, CNS, cardiovascular or gastrointestinal effects of the
XX toxicity.
PS Claim 14; Page 535; 735pp; English.
XX The present invention relates to novel GENSET polynucleotides (AB236404-
XX AB236911) encoding polypeptides (ABP75963-ABP76368). The polynucleotides
XX and polypeptides are useful in screening and diagnostic assays for
XX abnormal GENSET expression and/or biological activity. They are also
XX useful for screening of compounds for treating or preventing GENSET-
XX related disorders, such as heavy metal toxicity, cancer, inflammatory
XX diseases, immune disorders, and the neuromuscular, central nervous system
XX (CNS), cardiovascular or gastrointestinal effects of the toxicity
XX
SQ Sequence 65 AA;
Query Match 22.3%; Score 197; DB 6; Length 65;
Best Local Similarity 100.0%; Pred. No. 5.2e-09;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGSA 37
DB 25 DEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGSA 61
|||||
RESULT 14
AAR20165
ID AAR20165 standard; protein; 32 AA.
XX
AC AAR20165;
XX
XX 25-MAR-2003 (revised)
DT 01-APR-1992 (first entry)
XX
XX Apo AI polypeptide (B).
XX Monoclonal antibody; MAB AI-14; epitope; diagnosis.
KW
XX Synthetic.
XX
XX Key Location/Qualifiers
FT Peptide 1..32
FT Peptide 1..21
FT Peptide 5..28
FT Peptide 5..21
FT Peptide 6..28
FT Peptide 6..21
FT Misc-difference 11
FT /label= GLU, PHE
XX
XX WO9118619-A.

XX 12-DEC-1991.
PD
XX 07-JUN-1990; 90US-00534761.
XX
XX 07-JUN-1990; 90US-00534761.
XX (SCRI) SCRIPPS CLINIC & RES.
XX
XX Curtiss LK, Banka CL, Bonnet DJ, Smith RS;
PI WPI; 1992-007201/01.
XX
XX New Apo AI polypeptide(s) and anti-Apo antibodies - used to detect Apo AI
PT in vascular fluid samples and increase LCAT-mediated cholesterol
PT esterification in humans.
XX
XX Claim 5; Page 68; 87pp; English.
XX The peptides represented in AAR20164-65 are capable of immunologically
XX mimicking an Apo AI epitope. They are useful in diagnosis and detection
XX of Apo AI in fluids and for preparing anti-Apo AI antibodies. It is
XX useful in therapeutic methods for increasing LCAT-mediated cholesterol
XX esterification in humans. The Apo AI polypeptide is selected from the
XX peptides indicated in the features and includes amino acids 6-21 defining
XX a conserved native epitope on Apo AI capable of immunoreacting with
XX monoclonal antibody MAB AI-4. (Updated on 25-MAR-2003 to correct PA
XX field.)
SQ Sequence 32 AA;
Query Match 18.4%; Score 163; DB 2; Length 32;
Best Local Similarity 96.9%; Pred. No. 1.5e-06;
Matches 31; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 94 KAKQVPYLDLDFQKKQWQEMELYRQKVEPLRAE 125
DB 1 KAKQVPYLDLDFQKKQWQEMELYRQKVEPLRAE 32
|||||
RESULT 15
ABP31744
ID ABP31744 standard; protein; 64 AA.
XX
XX AC ABP31744;
XX
XX 08-JUL-2002 (first entry)
DT
XX Human apolipoprotein-like ORF717 protein, SEQ ID NO:1434.
DE
XX Human; ORF; open reading frame; ORF; drug screening; diagnosis;
KW disease monitoring; cytokine; cell proliferation; cell differentiation;
KW immune modulation; haematopoiesis regulation; tissue growth;
KW angiogenesis; activin; inhibin; chemotactic; chemokinetic; haemostatic;
KW thrombolytic; tumour inhibition; bodily characteristic; fertility;
KW behaviour; cancer; proliferative disorder; neurological disorder;
KW cardiovascular disease; immune system disorder; organ transplantation;
KW tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
KW hypothyroidism; cholesterol ester storage disease; infection; vulnery;
KW vasotropic; antipsoriatic; antidiabetic; cytostatic; nootropic;
KW neuroprotective; antiatherosclerotic; anticoagulant; thrombolytic;
KW cardiant; hypotensive; antithyroid; antinflammatory; immunomodulator;
KW dermatological; analgesic; virucide; antibacterial; fungicide.
XX
XX Homo sapiens.
OS
XX WO200190366-A2.
PN
XX 29-NOV-2001.
PD
XX 24-MAY-2001; 2001WO-US017076.
XX
XX 24-MAY-2000; 2000US-0206690P.
XX

XX (CURA-) CURAGEN CORP.
XX Leach MD, Shinkets RA;
XX WPI; 2002-106200/14.
XX N-PSDB; ABN75770.
XX Novel human polypeptides and polynucleotides useful for diagnosing,
PT preventing and treating cardiovascular disease, neurodegenerative,
PT hyperproliferative disorders and disorders related to organ
XX transplantation.
XX
PS Claim 10; Page 625; 2508pp; English.
XX
XX Sequences ABP31028-ABP35561 represent 4534 novel human proteins
CC designated ORF (open reading frame) 1-4534, and sequences ABN75054-
CC ABN75987 represent cDNAs encoding them. The invention also encompasses
CC polypeptides at least 80% identical to the ORF1-ORF4534 (collectively
CC referred to as ORFX) proteins, polynucleotides at least 85% identical to
CC the ORFX nucleic acid sequences, vectors and host cells comprising ORFX
CC polynucleotides, the recombinant production of ORFX proteins, antibodies
CC specific for ORFX proteins, methods of detecting ORFX polynucleotides and
CC polypeptides, methods of screening for modulators of ORFX expression or
CC activity, and methods of screening individuals for a predisposition to an
CC ORFX-associated disorder. The ORFX proteins of the invention have a wide
CC range of biological activities, such as cytokine, cell proliferation,
CC cell differentiation, immune modulation, haematopoiesis regulation,
CC tissue growth, angiogenesis, activin or inhibin activity, chemotactic/
CC chemokinetic activity, haemostatic activity, thrombolytic activity,
CC receptor/ligand, antiinflammatory activity, tumour inhibition activity,
CC and antiinfective activity, and may also be involved in the determination
CC of bodily characteristics, fertility and behaviour. ORFX proteins,
CC nucleic acids and antibodies may be used in the treatment of cancers,
CC other proliferative disorders such as psoriasis and benign tumours,
CC neurological disorders such as epilepsy and Alzheimer's disease,
CC cardiovascular diseases, immune system disorders, disorders related to
CC organ transplantation, disorders of tissue growth and regeneration,
CC diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester
CC storage disease, and infectious diseases caused by viral, bacterial,
CC fungal and other pathogens. ORFX nucleic acids may also be used as a
CC source of primers and probes, in the detection of ORFX genomic sequences
CC or transcripts, in the identification and cloning of homologous
CC sequences, in genetic diagnosis, and in forensic biology. The ORFX
CC nucleic acids may additionally be used to produce transgenic animals
CC which may be useful for studying the function and/or activity of ORFX
CC protein, and in drug screening. The ORFX proteins may also be used as
CC immunogens to generate specific antibodies, which are useful in the
CC diagnosis, treatment and monitoring of ORFX-associated diseases
XX
SQ Sequence 64 AA;

Query Match 18.3%; Score 162; DB 5; Length 64;
Best Local Similarity 75.0%; Pred. No. 4.2e-06;
Matches 30; Conservative 4; Mismatches 6; Indels 0; Caps 0;

QY 2 EPPQSPWDRVKDLATVVDVLKDSGRDYVSQFSGSALGKQ 41
: ||| ||||| ||||| : ||||| ||||| |||||
Db 25 DEPQSQWDRVKDPATVVDVAVKDSGRNYSQFESSTLGGQ 64

Search completed: December 21, 2004, 10:46:20
Job time : 69 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2004, 13:22:55 ; Search time 144 Seconds
(without alignments)
422.444 Million cell updates/sec

Title: US-09-803-918A-2_COPY_25_194

Perfect score: 885

Sequence: 1 DEPPQSPMDRVKDLATVYVD.....RARAHVDALRTHLAPYSDEL 170

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1589859 seqs, 357834939 residues

Total number of hits satisfying chosen parameters: 921198

Minimum DB seq length: 0

Maximum DB seq length: 170

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	885	100.0	170	9	US-09-803-918A-3
2	647.5	73.2	166	14	US-10-291-172-729
3	647.5	73.2	166	15	US-10-221-278-729
4	529.5	59.8	168	15	US-10-465-789A-29
5	529.5	59.8	168	15	US-10-465-789A-82
6	451.5	51.0	93	15	US-10-038-854-403
7	451.5	51.0	93	15	US-10-038-854-405
8	442.5	50.0	168	15	US-10-465-789A-23
9	437	49.4	168	15	US-10-408-765A-2201
10	396.5	44.8	168	15	US-10-465-789A-81
11	318	35.9	86	15	US-10-424-599-276814
12	227	25.6	43	15	US-10-465-789A-69
13	162	18.3	64	11	US-09-864-408A-1434

14	137	15.5	26	14	US-10-120-508-21	Sequence 21, Appl
15	122	13.8	22	15	US-10-465-789A-52	Sequence 52, Appl
16	118	13.3	22	15	US-10-465-789A-50	Sequence 50, Appl
17	113	12.8	22	15	US-10-465-789A-54	Sequence 54, Appl
18	112	12.7	22	15	US-10-465-789A-49	Sequence 49, Appl
19	108	12.2	22	15	US-10-465-789A-53	Sequence 53, Appl
20	106.5	12.0	107	14	US-10-023-066A-92	Sequence 92, Appl
21	96.5	10.9	95	17	US-10-487-096-2	Sequence 2, Appl
22	94	10.6	18	14	US-10-142-238A-10	Sequence 10, Appl
23	94	10.6	30	14	US-10-142-238A-40	Sequence 40, Appl
24	94	10.6	30	14	US-10-142-238A-41	Sequence 41, Appl
25	92	10.4	18	14	US-10-142-238A-33	Sequence 33, Appl
26	91	10.3	158	15	US-10-424-599-198597	Sequence 198597, Ap
27	90	10.2	18	14	US-10-142-238A-13	Sequence 13, Appl
28	90	10.2	160	9	US-09-764-864-1129	Sequence 1129, Ap
29	90	10.2	162	17	US-10-425-115-186898	Sequence 186898, Ap
30	89	10.1	16	17	US-10-700-340-37	Sequence 37, Appl
31	89	10.1	30	14	US-10-142-238A-39	Sequence 39, Appl
32	89	10.1	30	14	US-10-142-238A-43	Sequence 43, Appl
33	88	9.9	129	10	US-09-769-787-51	Sequence 51, Appl
34	87	9.8	17	15	US-10-601-100-44	Sequence 44, Appl
35	87	9.8	18	14	US-10-142-238A-32	Sequence 32, Appl
36	87	9.8	18	14	US-10-142-238A-35	Sequence 35, Appl
37	87	9.8	89	17	US-10-425-115-191061	Sequence 191061, Ap
38	86	9.7	18	14	US-10-142-238A-9	Sequence 9, Appl
39	86	9.7	18	14	US-10-142-238A-11	Sequence 11, Appl
40	86	9.7	18	14	US-10-142-238A-14	Sequence 14, Appl
41	86	9.7	30	14	US-10-142-238A-42	Sequence 42, Appl
42	86	9.7	30	14	US-10-142-238A-46	Sequence 46, Appl
43	85.5	9.7	109	14	US-10-029-386-29332	Sequence 29332, A
44	85.5	9.7	165	16	US-10-767-701-42732	Sequence 42732, A
45	85	9.6	18	14	US-10-142-238A-12	Sequence 12, Appl

ALIGNMENTS

RESULT 1

US-09-803-918A-3
; Sequence 3, Application US/09803918A
; Patent No. US20020064820A1
; GENERAL INFORMATION:
; APPLICANT: Dayer, Jean-Michel
; APPLICANT: Burger, Danielle
; APPLICANT: Kohno, Tadahiko
; APPLICANT: Edwards III, Carl K.
; TITLE OF INVENTION: APO-A-1 REGULATION OF T-CELL SIGNALING
; FILE REFERENCE: 06843.0035-00000
; CURRENT APPLICATION NUMBER: US/09/803,918A
; CURRENT FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: 60/189,008
; PRIOR FILING DATE: 2000-03-13
; PRIOR APPLICATION NUMBER: 60/193,551
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(170)
; OTHER INFORMATION: 18 kDa N-terminal fragment
US-09-803-918A-3

Query Match 100.0%; Score 885; DB 9; Length 170;

Best Local Similarity 100.0%; Pred. No. 1.9e-64;

Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DEPPQSPMDRVKDLATVYVDVLKDSGRDYVSQFEGSALGKQLNKLDDNWDVSTSTFSKL 60

DB 1 DEPPQSPMDRVKDLATVYVDVLKDSGRDYVSQFEGSALGKQLNKLDDNWDVSTSTFSKL 60

QY 61 REQLGPVTOEFWDNLEKETEGRLQEMSKDLEEVKAKVQPYLDDFQKKWOEEMELYRQKVE 120
Db 61 REQLGPVTOEFWDNLEKETEGRLQEMSKDLEEVKAKVQPYLDDFQKKWOEEMELYRQKVE 120
QY 121 PLRAELQEGARQKLHLEQLKSPLGEMDRARAHVDALRTHLAPYSDEL 170
Db 121 PLRAELQEGARQKLHLEQLKSPLGEMDRARAHVDALRTHLAPYSDEL 170

RESULT 2
US-10-291-172-729
; Sequence 729, Application US/10291172
; Publication No. US20030228584A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: No. US20030228584A1e1 Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-045
; CURRENT APPLICATION NUMBER: US/10/291,172
; CURRENT FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 09/693,267
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/665,363
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 09/616,847
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 09/596,193
; PRIOR FILING DATE: 2000-06-17
; PRIOR APPLICATION NUMBER: 09/574,454
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: 09/519,705
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 752
; SEQ ID NO 729
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-291-172-729

Query Match 73.2%; Score 647.5; DB 14; Length 166;
Best Local Similarity 82.7%; Pred. No. 4.4e-45;
Matches 129; Conservative 1; Mismatches 3; Indels 23; Gaps 1;
QY 1 DEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGSALGKQLNLKLLDNWDSVTFSKL 60
Db 34 DEPPQSPWDRVKDLATVYVDVLKDSGK-----DSVTSTFSKL 70
QY 61 REQLGPVTOEFWDNLEKETEGRLQEMSKDLEEVKAKVQPYLDDFQKKWOEEMELYRQKVE 120
Db 71 REQLGPVTOEFWDNLEKETEGRLQEMSKDLEEVKAKVQPYLDDFQKKWOEEMELYRQKVE 130
QY 121 PLRAELQEGARQKLHLEQLKSPLGEMDRARAHV 156
Db 131 PLRAELQEGARQKLHLEQLKSPLGEDADRRRAHV 166

RESULT 3
US-10-221-278-729
; Sequence 729, Application US/10221278
; Publication No. US20040034208A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: No. US20040034208A1e1 Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-045
; CURRENT APPLICATION NUMBER: US/10/221,278
; CURRENT FILING DATE: 2002-09-06
; PRIOR APPLICATION NUMBER: 09/693,267
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/665,363
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 09/616,847
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 09/596,193

; PRIOR FILING DATE: 2000-06-17
; PRIOR APPLICATION NUMBER: 09/574,454
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: 09/519,705
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 752
; SEQ ID NO 729
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-221-278-729
Query Match 73.2%; Score 647.5; DB 15; Length 166;
Best Local Similarity 82.7%; Pred. No. 4.4e-45;
Matches 129; Conservative 1; Mismatches 3; Indels 23; Gaps 1;
QY 1 DEPPQSPWDRVKDLATVYVDVLKDSGRDYSQFEGSALGKQLNLKLLDNWDSVTFSKL 60
Db 34 DEPPQSPWDRVKDLATVYVDVLKDSGK-----DSVTSTFSKL 70
QY 61 REQLGPVTOEFWDNLEKETEGRLQEMSKDLEEVKAKVQPYLDDFQKKWOEEMELYRQKVE 120
Db 71 REQLGPVTOEFWDNLEKETEGRLQEMSKDLEEVKAKVQPYLDDFQKKWOEEMELYRQKVE 130
QY 121 PLRAELQEGARQKLHLEQLKSPLGEMDRARAHV 156
Db 131 PLRAELQEGARQKLHLEQLKSPLGEDADRRRAHV 166

RESULT 4
US-10-465-789A-29
; Sequence 29, Application US/10465789A
; Publication No. US2004005384A1
; GENERAL INFORMATION:
; APPLICANT: Sligar, Stephen G
; APPLICANT: Bayburt, Timothy H
; APPLICANT: Schuler, Mary A
; APPLICANT: Civjan, Natanya R
; APPLICANT: Yelena V. Grinkova
; APPLICANT: Ilia G. Denisov
; TITLE OF INVENTION: Membrane Scaffold Proteins
; FILE REFERENCE: 87-00A
; CURRENT APPLICATION NUMBER: US/10/465,789A
; CURRENT FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: 09/990,087
; PRIOR FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 60/252,233
; PRIOR FILING DATE: 2000-11-20
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 29
; LENGTH: 168
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: His-tagged MSP1D6
US-10-465-789A-29

Query Match 59.8%; Score 529.5; DB 15; Length 168;
Best Local Similarity 74.5%; Pred. No. 1.9e-35;
Matches 108; Conservative 5; Mismatches 25; Indels 7; Gaps 1;
QY 26 GRDYSQFEGSALGKQLNLKLLDNWDSVTSTFSKLREQLGPVTOEFWDNLEKETEGRLQ 85
Db 2 GHHHHHHHTEG-----RLKLLDNWDSVTSTFSKLREQLGPVTOEFWDNLEKETEGRLQ 54
QY 86 MSKDLLEEVKAKVQPYLDDFQKKWOEEMELYRQKVEPLRAELQEGARQKLHLEQLKSPLG 145
Db 55 MSKDLLEEVKAKVQPYLDDFQKKWOEEMELYRQKVEPLRAELQEGARQKLHLEQLKSPLG 114
QY 146 EEMDRARAHVDALRTHLAPYSDEL 170
Db 115 AEYHAKATEHLSTLSEKAKPALEDL 139


```
RESULT 5
US-10-465-789A-82
; Sequence 82, Application US/10465789A
; Publication No. US20040053384A1
; GENERAL INFORMATION:
; APPLICANT: Sligar, Stephen G
; APPLICANT: Bayburt, Timothy H
; APPLICANT: Schuler, Mary A
; APPLICANT: Civjan, Natanya R
; APPLICANT: Yelena V. Grinkova
; APPLICANT: Ilija G. Denisov
; TITLE OF INVENTION: Membrane Scaffold Proteins
; FILE REFERENCE: 87-00A
; CURRENT APPLICATION NUMBER: US/10/465,789A
; CURRENT FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: 09/990,087
; PRIOR FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 60/252,233
; PRIOR FILING DATE: 2000-11-20
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 82
; LENGTH: 168
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: His-tagged MSP1d6D7
US-10-465-789A-82

Query Match          59.8%; Score 529.5; DB 15; Length 168;
Best Local Similarity 74.5%; Pred. No. 1.9e-35;
Matches 108; Conservative 5; Mismatches 25; Indels 7; Gaps 1;

QY 26 GRDYSQFGSALGKQLNLKLLDNWDSVTSTFSKLREQLGPTQEFWDNLEKETEGLRQE 85
DB 2 GHHHHHHIEG-----RLKLLDNWDSVTSTFSKLREQLGPTQEFWDNLEKETEGLRQE 54

QY 86 MSKDLSEVAKVQPYLLDDFQKKWQEMELYRQKVEPLRAELQEGARQKLHEQKLSPLG 145
DB 55 MSKDLSEVAKVQPYLLDDFQKKWQEMELYRQKVEPLRAELQEGARQKLHEQKLSARL 114

QY 146 EEMDRARAHVDALRTHLAPYSDEL 170
DB 115 AEVHAKATEHLSTLSEKAKPALEDL 139

RESULT 6
US-10-038-854-403
; Sequence 403, Application US/10038854
; Publication No. US20040022781A1
; GENERAL INFORMATION:
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Li, Li
; APPLICANT: Wolenc, Adam R
; APPLICANT: Vernet, Corine
; APPLICANT: Eisen, Andrew J
; APPLICANT: Liu, Xiaohong
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Shimkets, Richard A
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spaderna, Steven K
; APPLICANT: Gorman, Linda
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Gangolli, Beha A
; APPLICANT: Guo, Xiaojia S
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Rastelli, Luca
; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc
```

```
; APPLICANT: Burgess, Catherine E
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gunther, Erik
; APPLICANT: Smithson, Glenda
; APPLICANT: Millet, Isabelle
; APPLICANT: MacDougall, John R
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-230
; CURRENT APPLICATION NUMBER: US/10/038,854
; CURRENT FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/258,928
; PRIOR FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: 60/259,415
; PRIOR FILING DATE: 2001-01-02
; PRIOR APPLICATION NUMBER: 60/259,785
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/269,814
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/279,832
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,833
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,863
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/283,889
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,447
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/286,683
; PRIOR FILING DATE: 2001-04-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 403
; LENGTH: 93
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-038-854-403

Query Match          51.0%; Score 451.5; DB 15; Length 93;
Best Local Similarity 77.5%; Pred. No. 2.1e-29;
Matches 93; Conservative 0; Mismatches 0; Indels 27; Gaps 1;

QY 14 LATVYVDVLKDSGRDYSQFGSALGKQLNLKLLDNWDSVTSTFSKLREQLGPTQEFWD 73
DB 1 LATVYVDVLK-----DSVTSTFSKLREQLGPTQEFWD 33

QY 74 NLEKETEGLRQEMSKDLSEVAKVQPYLLDDFQKKWQEMELYRQKVEPLRAELQEGARQK 133
DB 34 NLEKETEGLRQEMSKDLSEVAKVQPYLLDDFQKKWQEMELYRQKVEPLRAELQEGARQK 93

RESULT 7
US-10-038-854-405
; Sequence 405, Application US/10038854
; Publication No. US20040022781A1
; GENERAL INFORMATION:
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Li, Li
; APPLICANT: Wolenc, Adam R
; APPLICANT: Vernet, Corine
; APPLICANT: Eisen, Andrew J
; APPLICANT: Liu, Xiaohong
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Shimkets, Richard A
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spaderna, Steven K
; APPLICANT: Gorman, Linda
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Gangolli, Esha A
```

```
; APPLICANT: Guo, Xiaojia S
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Rastelli, Luca J
; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc
; APPLICANT: Burgess, Catherine E
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gunther, Erik
; APPLICANT: Smithson, Glenda
; APPLICANT: Millet, Isabelle
; APPLICANT: MacDougall, John R
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-230
; CURRENT APPLICATION NUMBER: US/10/038,854
; CURRENT FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/258,928
; PRIOR FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: 60/259,415
; PRIOR FILING DATE: 2001-01-02
; PRIOR APPLICATION NUMBER: 60/259,785
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/269,814
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/279,832
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,833
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,863
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/283,889
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,447
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/286,683
; PRIOR FILING DATE: 2001-04-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 405
; TYPE: PRT
; LENGTH: 93
; ORGANISM: Homo sapiens
US-10-038-854-405

Query Match 51.0%; Score 451.5; DB 15; Length 93;
Best Local Similarity 77.5%; Pred. No. 2.1e-29;
Matches 93; Conservative 0; Mismatches 0; Indels 27; Gaps 1;

QY 14 LATVYDVVLKDSGRDYVSQFEGSALGKQLNLKLLDNWDSVTSTFSKLRQLGVPVTOEFWDNLEKETEGLRQE 73
Db 1 LATVYDVVLK-----DSVTSTFSKLRQLGVPVTOEFWDNLEKETEGLRQE 33

QY 74 NLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQOEEMELYRQKVEPLRAELQEGAROK 133
Db 34 NLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQOEEMELYRQKVEPLRAELQEGAROK 93

RESULT 8
US-10-465-789A-23
; Sequence 23, Application US/10465789A
; Publication No. US20040053384A1
; GENERAL INFORMATION:
; APPLICANT: Sligar, Stephen G
; APPLICANT: Bayburt, Timothy H
; APPLICANT: Schuler, Mary A
; APPLICANT: Civjan, Natanva R
; APPLICANT: Yelena V. Grinkova
; APPLICANT: Ilia G. Denisov
; TITLE OF INVENTION: Membrane Scaffold Proteins
; FILE REFERENCE: 87-00A
; CURRENT APPLICATION NUMBER: US/10/465,789A
; CURRENT FILING DATE: 2003-06-18
```

```
; PRIOR APPLICATION NUMBER: 09/990,087
; PRIOR FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 60/252,233
; PRIOR FILING DATE: 2000-11-20
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 23
; LENGTH: 168
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: His-tagged MSP1D5D6
US-10-465-789A-23

Query Match 50.0%; Score 442.5; DB 15; Length 168;
Best Local Similarity 63.4%; Pred. No. 2.4e-28;
Matches 92; Conservative 9; Mismatches 37; Indels 7; Gaps 1;

QY 26 GRDYVSQFEGSALGKQLNLKLLDNWDSVTSTFSKLRQLGVPVTOEFWDNLEKETEGLRQE 85
Db 2 GHHHHHHIEG-----RLKLLDNWDSVTSTFSKLRQLGVPVTOEFWDNLEKETEGLRQE 54

QY 86 MSKDL EEVKAKVQPYLDDFQKKWQOEEMELYRQKVEPLRAELQEGAROKLHLOEKLSPLG 145
Db 55 MSKDL EEVKAKVQPYLDDFQKKWQOEEMELYRQKVEPLRAELQEGAROKLHLOEKLSPLG 114

QY 146 EEMRDRARAHVDALRTHLAPYSDEL 170
Db 115 AEYHAKATEHLSTLSEKAKPALEDL 139

RESULT 9
US-10-408-765A-2201
; Sequence 2201, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2201
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-2201

Query Match 49.4%; Score 437; DB 16; Length 158;
Best Local Similarity 100.0%; Pred. No. 6.2e-28;
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 MSKDL EEVKAKVQPYLDDFQKKWQOEEMELYRQKVEPLRAELQEGAROKLHLOEKLSPLG 145
Db 1 MSKDL EEVKAKVQPYLDDFQKKWQOEEMELYRQKVEPLRAELQEGAROKLHLOEKLSPLG 60

QY 146 EEMRDRARAHVDALRTHLAPYSDEL 170
Db 61 EEMRDRARAHVDALRTHLAPYSDEL 85

RESULT 10
US-10-465-789A-81
; Sequence 81, Application US/10465789A
; Publication No. US20040053384A1
```

```

; GENERAL INFORMATION:
; APPLICANT: Sligar, Stephen G
; APPLICANT: Bayburt, Timothy H
; APPLICANT: Schuler, Mary A
; APPLICANT: Civjan, Natanva R
; APPLICANT: Yelena V. Grinkova
; APPLICANT: Ilia G. Denisov
; TITLE OF INVENTION: Membrane Scaffold Proteins
; FILE REFERENCE: 87-00A
; CURRENT APPLICATION NUMBER: US/10/465,789A
; PRIOR FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: 09/990,087
; PRIOR FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 60/252,233
; PRIOR FILING DATE: 2000-11-20
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 81
; LENGTH: 168
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: artificial sequence of MSPID4D5
US-10-465-789A-81

Query Match      44.8%; Score 396.5; DB 15; Length 168;
Best Local Similarity 65.4%; Pred. No. 1.4e-24;
Matches 83; Conservative 0; Mismatches 0; Indels 44; Gaps 1;

QY 44 LKLLDNWDSVTFSKLRQLGPGVTQEFWDNLEKETEGRLQEMSKDLSEVKAKVQPYLDD 103
Db 13 LKLLDNWDSVTFSKLRQLGPGVTQEFWDNLEKETEGRLQEMSKDLSEVKAKVQ----- 67

QY 104 FQKKQOEMELYRQKVEPLRAELQEGARQKLHELQKLSPLGEMRDRARAHVDALRTHL 163
Db 68 -----PLGEMRDRARAHVDALRTHL 88

QY 164 APYSDEL 170
Db 89 APYSDEL 95

RESULT 11
US-10-424-599-276814
; Sequence 276814, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(33223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; PRIOR FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 276814
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_91983C.1.pgp
US-10-424-599-276814

Query Match      35.9%; Score 318; DB 15; Length 86;
Best Local Similarity 98.4%; Pred. No. 1.5e-18;
Matches 61; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DEPPQSPWDRVKDLATVYVDVLKDSGRDYVSQFEGSALGKQLNLKLLDNWDSVTFSKL 60
Db 25 DEPPQSPWDRVKDLATVYVDVLKDSGRDYVSQFEGSALGKQLNLKLLDNWDSVTFSKL 84
```

```

QY 61 RE 62
Db 85 RE 86

RESULT 12
US-10-465-789A-89
; Sequence 89, Application US/10465789A
; Publication No. US20040053384A1
; GENERAL INFORMATION:
; APPLICANT: Sligar, Stephen G
; APPLICANT: Bayburt, Timothy H
; APPLICANT: Schuler, Mary A
; APPLICANT: Civjan, Natanva R
; APPLICANT: Yelena V. Grinkova
; APPLICANT: Ilia G. Denisov
; TITLE OF INVENTION: Membrane Scaffold Proteins
; FILE REFERENCE: 87-00A
; CURRENT APPLICATION NUMBER: US/10/465,789A
; PRIOR FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: 09/990,087
; PRIOR FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 60/252,233
; PRIOR FILING DATE: 2000-11-20
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 89
; LENGTH: 43
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: sequence of GLOB
US-10-465-789A-89

Query Match      25.6%; Score 227; DB 15; Length 43;
Best Local Similarity 100.0%; Pred. No. 1.8e-11;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DEPPQSPWDRVKDLATVYVDVLKDSGRDYVSQFEGSALGKQLN 43
Db 1 DEPPQSPWDRVKDLATVYVDVLKDSGRDYVSQFEGSALGKQLN 43

RESULT 13
US-09-864-408A-1434
; Sequence 1434, Application US/09864408A
; Publication No. US20040009474A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Shimkets, Richard A.
; TITLE OF INVENTION: No. US20040009474A1el Human Polynucleotides and Polypeptides Encr
; FILE REFERENCE: 21402-012
; CURRENT APPLICATION NUMBER: US/09/864,408A
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/206,690
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 9068
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1434
; LENGTH: 64
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-864-408A-1434

Query Match      18.3%; Score 162; DB 11; Length 64;
Best Local Similarity 75.0%; Pred. No. 5.7e-06;
Matches 30; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 2 BPQPSPWDRVKDLATVYVDVLKDSGRDYVSQFEGSALGKQ 41
Db 25 DEPPQSPWDRVKDLATVYVDVDAVDSGRNYVSQFESSTLGGQ 64
```

RESULT 14
US-10-120-508-21
; Sequence 21, Application US/10120508
; Publication No. US20030191057A1
; GENERAL INFORMATION:
; APPLICANT: FOGELMAN, ALAN
; APPLICANT: NAVAB, MOHAMAD
; TITLE OF INVENTION: G-TYPE PEPTIDES TO AMELIORATE ATHEROSCLEROSIS
; FILE REFERENCE: 407T-301100US
; CURRENT APPLICATION NUMBER: US/10/120,508
; CURRENT FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 21
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic D peptide.
US-10-120-508-21

Query Match 15.5%; Score 137; DB 14; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.00021;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 WDRVKDLATVYVDVLKDSGRDYVSQF 33
|||||
Db 1 WDRVKDLATVYVDVLKDSGRDYVSQF 26

RESULT 15
US-10-465-789A-52
; Sequence 52, Application US/10465789A
; Publication No. US20040053384A1
; GENERAL INFORMATION:
; APPLICANT: Sligar, Stephen G
; APPLICANT: Bayburt, Timothy H
; APPLICANT: Schuler, Mary A
; APPLICANT: Civjan, Nacanya R
; APPLICANT: Yelena V. Grinkova
; APPLICANT: Ilia G. Denisov
; TITLE OF INVENTION: Membrane Scaffold Proteins
; FILE REFERENCE: 87-00A
; CURRENT APPLICATION NUMBER: US/10/465,789A
; CURRENT FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: 09/990,087
; PRIOR FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 60/252,233
; PRIOR FILING DATE: 2000-11-20
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 52
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: artificial sequence of Helix 4
US-10-465-789A-52

Query Match 13.8%; Score 122; DB 15; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.0029;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 99 PYLDDFQKKQEMELYRQKVE 120
|||||
Db 1 PYLDDFQKKQEMELYRQKVE 22

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2004, 13:37:37 ; Search time 39 Seconds

(without alignments)
289.078 Million cell updates/sec

Title: US-09-803-918A-2_COPY_25_194

Perfect score: 885

Sequence: 1 DEPPQSPMDRVKDLATVYVD.....RAAHVDAALRTHLAPYSDEL 170

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 358685

Minimum DB seq length: 0
Maximum DB seq length: 170

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/prodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/prodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/prodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/prodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/prodata/1/iaa/PCTUS_COMB.pep.*
- 6: /cgn2_6/prodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	328	37.1	64	US-08-292-870-1	Sequence 1, Appli
2	227	25.6	44	US-08-292-870-2	Sequence 2, Appli
3	163	18.4	32	US-08-292-870-3	Sequence 3, Appli
4	124	14.0	25	US-08-292-870-4	Sequence 4, Appli
5	106.5	12.0	107	US-08-182-175A-105	Sequence 105, App
6	106.5	12.0	107	US-08-474-633A-92	Sequence 92, Appl
7	106.5	12.0	107	US-08-823-771-92	Sequence 92, Appl
8	106.5	12.0	107	PCT-US92-06412-105	Sequence 105, App
9	101	11.4	128	US-09-621-976-5220	Sequence 5220, Ap
10	100	11.3	110	US-07-849-389-7	Sequence 7, Appli
11	91	10.3	16	US-07-959-946-5	Sequence 5, Appli
12	91	10.3	16	US-08-333-577-5	Sequence 5, Appli
13	91	10.3	16	PCT-US92-08634-5	Sequence 5, Appli
14	89	10.1	114	US-09-079-030-129	Sequence 129, App
15	88	9.9	129	US-09-583-110-3496	Sequence 3496, Ap
16	84.5	9.5	105	US-08-743-200-2	Sequence 2, Appli
17	81.5	9.2	131	US-09-513-999C-5882	Sequence 5882, Ap
18	81	9.2	103	US-09-134-000C-4548	Sequence 4548, Ap
19	81	9.2	147	US-08-529-055-18	Sequence 18, Appl
20	81	9.2	153	US-08-529-055-23	Sequence 23, Appl
21	81	9.2	156	US-09-270-767-32734	Sequence 32734, A
22	81	9.2	156	US-09-270-767-47951	Sequence 47951, A
23	81	9.2	160	US-08-529-055-33	Sequence 33, Appl
24	80.5	9.1	159	US-09-248-796A-16325	Sequence 16325, A
25	80	9.0	72	US-08-182-175A-87	Sequence 87, Appl
26	80	9.0	72	PCT-US92-06412-87	Sequence 87, Appl
27	78	8.8	166	US-08-365-103B-14	Sequence 14, Appl

28	78	8.8	167	1	US-08-365-103B-12	Sequence 12, Appl
29	77.5	8.8	144	4	US-09-270-767-62028	Sequence 62028, A
30	77.5	8.8	164	4	US-09-270-767-59203	Sequence 59203, A
31	76.5	8.6	114	4	US-09-513-999C-6218	Sequence 6218, Ap
32	76	8.6	96	4	US-09-513-999C-5524	Sequence 5524, Ap
33	76	8.6	114	4	US-09-252-991A-18411	Sequence 18411, A
34	76	8.6	137	4	US-09-513-999C-7081	Sequence 7081, Ap
35	74.5	8.4	105	4	US-09-513-999C-6681	Sequence 6681, Ap
36	74.5	8.4	157	4	US-09-270-767-41567	Sequence 41567, A
37	73.5	8.3	99	4	US-09-270-767-44037	Sequence 44037, A
38	73	8.2	144	4	US-09-270-767-34931	Sequence 34931, A
39	73	8.2	144	4	US-09-270-767-50148	Sequence 50148, A
40	73	8.2	145	2	US-08-686-599A-20	Sequence 20, Appl
41	73	8.2	153	4	US-09-134-000C-5741	Sequence 5741, Ap
42	73	8.2	165	4	US-09-898-554-24	Sequence 24, Appl
43	73	8.2	166	4	US-09-270-767-59533	Sequence 59533, A
44	72	8.1	124	4	US-09-270-767-45785	Sequence 45785, A
45	71.5	8.1	108	2	US-08-710-749-24	Sequence 24, Appl

ALIGNMENTS

RESULT 1
US-08-292-870-1
; Sequence 1, Application US/08292870
; Patent No. 5814467
; GENERAL INFORMATION:
; APPLICANT: Curtiss, Linda K
; APPLICANT: Banks, Carole L
; APPLICANT: Bonnet, David J
; APPLICANT: Smith, Richard S
; TITLE OF INVENTION: APO AI POLYPEPTIDES, DIAGNOSTIC METHODS
; TITLE OF INVENTION: AND SYSTEMS FOR QUANTIFYING APO AI, AND THERAPEUTIC
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of
; ADDRESSEE: Patent Counsel
; STREET: 10666 No. 5814467th Torrey Pines Road., TPC 8
; CITY: La Jolla
; STATE: California
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/292,870
; FILING DATE: 17-AUG-1994
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/534,761
; FILING DATE: 07-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/711,333
; FILING DATE: 06-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US 91/04038
; FILING DATE: 07-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: TSRI 210.1 D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 64 amino acids
; TYPE: amino acid

```
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-08-292-870-1

Query Match      37.1%; Score 328; DB 2; Length 64;
Best Local Similarity 100.0%; Pred. No. 1.4e-22;
Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 85 ENSKOLEEVKAKVPYLDQFKKQWQEMELYRQKVEPLRAELQEGARQKLHELQKLSPL 144
   |||||
Db 1 ENSKOLEEVKAKVPYLDQFKKQWQEMELYRQKVEPLRAELQEGARQKLHELQKLSPL 60
   |||||

QY 145 GEEM 148
   |||||
Db 61 GEEM 64
   |||||

RESULT 2
US-08-292-870-2
; Sequence 2, Application US/08292870
; Patent No. 5814467
; GENERAL INFORMATION:
; APPLICANT: Curtiss, Linda K
; APPLICANT: Banks, Carole L
; APPLICANT: Bonnet, David J
; APPLICANT: Smith, Richard S
; TITLE OF INVENTION: APO AI POLYPEPTIDES, DIAGNOSTIC METHODS
; TITLE OF INVENTION: AND SYSTEMS FOR QUANTIFYING APO AI, AND THERAPEUTIC
; TITLE OF INVENTION: METHODS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of
; ADDRESSEE: Patent Counsel
; STREET: 10666 No. 5814467th Torrey Pines Road., TPC 8
; CITY: La Jolla
; STATE: California
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/292,870
; FILING DATE: 17-AUG-1994
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/534,761
; FILING DATE: 07-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/711,333
; FILING DATE: 06-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US 91/04038
; FILING DATE: 07-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: TSRI 210.1 D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 44 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-08-292-870-2

Query Match      18.4%; Score 163; DB 2; Length 32;
Best Local Similarity 96.9%; Pred. NO. 3.9e-08;
Matches 31; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-08-292-870-1

Query Match      25.6%; Score 227; DB 2; Length 44;
Best Local Similarity 100.0%; Pred. No. 1e-13;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 RQEMSKOLEEVKAKVPYLDQFKKQWQEMELYRQKVEPLRAEL 126
   |||||
Db 1 RQEMSKOLEEVKAKVPYLDQFKKQWQEMELYRQKVEPLRAEL 44
   |||||

RESULT 3
US-08-292-870-3
; Sequence 3, Application US/08292870
; Patent No. 5814467
; GENERAL INFORMATION:
; APPLICANT: Curtiss, Linda K
; APPLICANT: Banks, Carole L
; APPLICANT: Bonnet, David J
; APPLICANT: Smith, Richard S
; TITLE OF INVENTION: APO AI POLYPEPTIDES, DIAGNOSTIC METHODS
; TITLE OF INVENTION: AND SYSTEMS FOR QUANTIFYING APO AI, AND THERAPEUTIC
; TITLE OF INVENTION: METHODS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of
; ADDRESSEE: Patent Counsel
; STREET: 10666 No. 5814467th Torrey Pines Road., TPC 8
; CITY: La Jolla
; STATE: California
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/292,870
; FILING DATE: 17-AUG-1994
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/534,761
; FILING DATE: 07-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/711,333
; FILING DATE: 06-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US 91/04038
; FILING DATE: 07-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: TSRI 210.1 D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 32 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Region
; LOCATION: 11
; OTHER INFORMATION: /note= "Xaa can be either E (Glu)
; OTHER INFORMATION: or F (Phe)"
US-08-292-870-3

Query Match      18.4%; Score 163; DB 2; Length 32;
Best Local Similarity 96.9%; Pred. NO. 3.9e-08;
Matches 31; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

QY 94 KAKVQPYLDDFOKKQOEEMELTRQKVEPLRAE 125
|||||
Db 1 KAKVQPYLDDXQKKWOEEMELTRQKVEPLRAE 32
|||||

RESULT 4

US-08-292-870-4

; Sequence 4, Application US/08292870

; Patent No. 5814467

; GENERAL INFORMATION:

; APPLICANT: Curtiss, Linda K

; APPLICANT: Banks, Carole L

; APPLICANT: Bonnet, David J

; APPLICANT: Smith, Richard S

; TITLE OF INVENTION: APO AI POLYPEPTIDES, DIAGNOSTIC METHODS

; TITLE OF INVENTION: AND SYSTEMS FOR QUANTIFYING APO AI, AND THERAPEUTIC

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: The Scripps Research Institute, Office of

; ADDRESSEE: Patent Counsel

; STREET: 10666 No. 5814467th Torrey Pines Road., TPC 8

; CITY: La Jolla

; STATE: California

; COUNTRY: US

; ZIP: 92037

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/292,870

; FILING DATE: 17-AUG-1994

; CLASSIFICATION: 436

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/534,761

; FILING DATE: 07-JUN-1990

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/711,333

; FILING DATE: 06-JUN-1991

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/US 91/04038

; FILING DATE: 07-JUN-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: Fitting, Thomas

; REGISTRATION NUMBER: 34,163

; REFERENCE/DOCKET NUMBER: TSRI 210.1 D1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 619-554-2937

; TELEFAX: 619-554-6312

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 25 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; FRAGMENT TYPE: internal

US-08-292-870-4

Query Match

Best Local Similarity 14.0%; Score 124; DB 2; Length 25;

Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 87 SKDLEEVKAKVQPYLDDFOKKQOE 111

Db 1 SKDLEEVKAKVQPYLDDFOKKQOE 25

RESULT 5

US-08-182-175A-105

; Sequence 105, Application US/08182175A

; Patent No. 5559223

; GENERAL INFORMATION:

; APPLICANT: Saverio Carl Falco

; APPLICANT: Sharon J. Keeler

; APPLICANT: Janet A. Rice

; TITLE OF INVENTION: Synthetic Storage Proteins with Defined Structure Containing

; NUMBER OF SEQUENCES: 113

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: E.I. du Pont de Nemours and Company

; STREET: 1007 Market Street

; CITY: Wilmington

; STATE: Delaware

; COUNTRY: USA

; ZIP: 19898

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy Disk

; COMPUTER: Macintosh

; OPERATING SYSTEM: Macintosh System, 6.0

; SOFTWARE: Microsoft Word, 4.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/182,175A

; FILING DATE:

; CLASSIFICATION: 800

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/743,006

; FILING DATE: 9 August 1991

; ATTORNEY/AGENT INFORMATION:

; NAME: Linda Axamethy Floyd

; REGISTRATION NUMBER: 33,692

; REFERENCE/DOCKET NUMBER: BB-1031

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (302) 992-4929

; TELEFAX: (302) 892-7949

; TELEX: 835420

; INFORMATION FOR SEQ ID NO: 105:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 107 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-182-175A-105

Query Match

Best Local Similarity 12.0%; Score 106.5; DB 1; Length 107;

Matches 27; Conservative 28; Mismatches 35; Indels 9; Gaps 3;

QY 59 KLREOLGPVTQEFW-----DNLEKETGLRQEMSKDLSEVKA--KVQPYLDDFOKKQOE 110

Db 7 KLUKEEMAKMKDEMVKLKEEMKKLEEMKMKVMEEMKKLEEMKAMEDKMKWLEEMKKLEEMK 66

QY 111 EMELYRQKVEPLRAELQEGARQKLHELOFKLSPLGEEMR 149

Db 67 KMKVMEEMKKLE-EKMKAMEDKMKWLEEMKKLEEMK 104

RESULT 6

US-08-474-633A-92

; Sequence 92, Application US/08474633A

; Patent No. 5773691

; GENERAL INFORMATION:

; APPLICANT: E. I. DU PONT DE NEMOURS AND

; APPLICANT: COMPANY

; TITLE OF INVENTION: CHIMERIC GENES AND

; TITLE OF INVENTION: METHODS FOR INCREASING

; TITLE OF INVENTION: INCREASING THE LYSINE

; TITLE OF INVENTION: AND THREONINE CONTENT

; NUMBER OF SEQUENCES: 107

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: E. I. DU PONT DE NEMOURS

; ADDRESSEE: AND COMPANY

; STREET: 1007 MARKET STREET

; CITY: WILMINGTON

; STATE: DELAWARE

```

; COUNTRY: U.S.A.
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MICROSOFT WORD VERSION 2.0C
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,633A
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: BARBARA C. SIEGELL
; REGISTRATION NUMBER: 30,684
; REFERENCE/DOCKET NUMBER: BB-1037-C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-992-4931
; TELEX: 835420
; INFORMATION FOR SEQ ID NO: 92:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 107 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-474-633A-92

Query Match 12.0%; Score 106.5; DB 1; Length 107;
Best Local Similarity 27.3%; Pred. No. 0.02;
Matches 27; Conservative 28; Mismatches 35; Indels 9; Gaps 3;

QY 59 KLREQLGPTVQBFW-----DNLEKETGLRQEMSKDLEEVKA--KVQPYLDDFOKKWQE 110
Db 7 KLEEMAKMKDEMVKLEEMKKLEEMKVMEEKMKLEEMKAMEDKMKWLEEKMKLEEK 66

QY 111 EMELYRQKVEPLRAELQEGARQKLHELQKLSPLGEEMR 149
Db 67 KMKWEEKMKKLE-EQKAMEDKMKWLEEKMKLEEKMK 104

RESULT 7
US-08-823-771-92
; Sequence 92, Application US/08823771
; Patent No. 6459019
; GENERAL INFORMATION:
; APPLICANT: E. I. DU PONT DE NEMOURS AND COMPANY
; TITLE OF INVENTION: CHIMERIC GENES AND METHODS FOR INCREASING THE LYSINE AND THREONINE CONTENT
; NUMBER OF SEQUENCES: 107
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: U.S.A.
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MICROSOFT WORD VERSION 2.0C
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/823,771
; FILING DATE: 24-Mar-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/474,633
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:

```

```

; NAME: BARBARA C. SIEGELL
; REGISTRATION NUMBER: 30,684
; REFERENCE/DOCKET NUMBER: BB-1037-C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-992-4931
; TELEX: 835420
; INFORMATION FOR SEQ ID NO: 92:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 107 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 92:
US-08-823-771-92

Query Match 12.0%; Score 106.5; DB 4; Length 107;
Best Local Similarity 27.3%; Pred. No. 0.02;
Matches 27; Conservative 28; Mismatches 35; Indels 9; Gaps 3;

QY 59 KLREQLGPTVQBFW-----DNLEKETGLRQEMSKDLEEVKA--KVQPYLDDFOKKWQE 110
Db 7 KLEEMAKMKDEMVKLEEMKKLEEMKVMEEKMKLEEMKAMEDKMKWLEEKMKLEEK 66

QY 111 EMELYRQKVEPLRAELQEGARQKLHELQKLSPLGEEMR 149
Db 67 KMKWEEKMKKLE-EQKAMEDKMKWLEEKMKLEEKMK 104

RESULT 8
PCT-US92-06412-105
; Sequence 105, Application PC/TUS9206412
; GENERAL INFORMATION:
; APPLICANT: Saverio Carl Falco
; APPLICANT: Sharon J. Keeler
; APPLICANT: Janet A. Rice
; TITLE OF INVENTION: Synthetic Storage Proteins with Defined Structure Containing P
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E.I. du Pont de Nemours and Company
; STREET: 1007 Market Street
; CITY: Wilmington
; STATE: Delaware
; COUNTRY: USA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: Macintosh System, 6.0
; SOFTWARE: Microsoft Word, 4.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/06412
; FILING DATE: 19920807
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/743,006
; FILING DATE: 9 August 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Linda Axamethy Floyd
; REGISTRATION NUMBER: 33,692
; REFERENCE/DOCKET NUMBER: BB-1031
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (302) 992-4929
; TELEX: 835420
; INFORMATION FOR SEQ ID NO: 105:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 107 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US92-06412-105

```


Query Match 12.0%; Score 106.5; DB 5; Length 107;
Best Local Similarity 27.3%; Pred. No. 0.02;
Matches 27; Conservative 28; Mismatches 35; Indels 9; Gaps 3;

QY 59 KLREQLGPTVQFPM-----DNLEKETEGLRQEMSKDLSEVKA--KVQPYLDDFQKKWQE 110
DB 7 KLKEENAKMKDEMMKLKEEMKLEERKMVBEERKKLEKMKAMEDKMKWLEEMKMKLEB 66

QY 111 EMELYRQKVEPLRAELQEGARQKLHELOEKLSPLGSEMR 149

DB 67 KMKVMEKKKLEB-EKMKAMEDKMKWLEERKMKLEEMK 104

RESULT 9

US-09-621-976-5220
Sequence 5220, Application US/09621976

Patent No. 6639063

GENERAL INFORMATION:

APPLICANT: Dumas Milne Edwards, J.B.

APPLICANT: Giordano, J.Y.

TITLE OF INVENTION: ESTs and Encoded Human Proteins.

FILE REFERENCE: GENSET 054PR2

CURRENT APPLICATION NUMBER: US/09/621,976

CURRENT FILING DATE: 2000-07-21

NUMBER OF SEQ ID NOS: 19335

SOFTWARE: Patent.pm

SEQ ID NO 5220

LENGTH: 128

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: SIGNAL

LOCATION: -23...-1

NAME/KEY: UNSURE

LOCATION: 104

OTHER INFORMATION: Xaa = His,Gln

US-09-621-976-5220

Query Match 11.4%; Score 101; DB 4; Length 128;

Best Local Similarity 29.0%; Pred. No. 0.08;
Matches 29; Conservative 22; Mismatches 39; Indels 10; Gaps 3;

QY 24 DSGRDYVSFEFGALGQKLKLLDNWDSVTFSKLREQLGPTVQFWDNLEKETEGLR 83

DB 37 DKGR--VEQIHQQMKARE-PATLKDSLEQDLNNMKFLEKRLPLSGSEAPRLPQDPVGR 93

QY 84 QEMSKLEEVKAKVQPYLDDFQKKWQEMELYRQKVEPLR 123

DB 94 RLQEELEEVKARLPQYM-----AEAHVLGVNLEGLR 126

RESULT 10

US-07-849-389-7

Sequence 7, Application US/07849389

Patent No. 5525493

GENERAL INFORMATION:

APPLICANT: HORNES, Erik

APPLICANT: UHLEN, Mathias

TITLE OF INVENTION: CLONING METHOD AND KIT

NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:

ADDRESSEE: Foley & Lardner

STREET: 1800 Diagonal Road, Suite 500

CITY: Alexandria

STATE: Virginia

COUNTRY: USA

ZIP: 22313-0299

COMPUTER READABLE FORM:

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/849,389

FILING DATE: 19920519

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: BENT, Stephen A.

REGISTRATION NUMBER: 29,768

REFERENCE/DOCKET NUMBER: 16787/168/DFBC

TELEPHONE: (703)836-9300

TELEFAX: (703)683-4109

TELEX: 899149

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 110 amino acids

TYPE: AMINO ACID

TOPOLOGY: linear

MOLECULE TYPE: protein

US-07-849-389-7

Query Match 11.3%; Score 100; DB 1; Length 110;

Best Local Similarity 26.6%; Pred. No. 0.081;

Matches 29; Conservative 16; Mismatches 42; Indels 22; Gaps 1;

QY 58 SKLREQLGPTVQFWDNLEKETEGLRQEMSKDLSEVKAQVQPYLDDFQKKWQEMELYRQ 117

DB 8 SELEQLTPVAETRAELSKELQAAEAPLGADMEDVGRGLVQY----- 50

QY 118 KVEPLRAELQEGARQKLHELOEKLSPLGSEMRDRARAHVDALRTHLAPY 166

DB 51 -----RGEVQAMLGQSTELRVRLASHURKRLRLRDADDLQKRLAVY 94

RESULT 11

US-07-959-946-5

Sequence 5, Application US/07959946

Patent No. 5408038

GENERAL INFORMATION:

APPLICANT: Smith, Richard K.

APPLICANT: Koduri, Raju

APPLICANT: Young, Stephen G.

APPLICANT: Witztum, Joseph L.

APPLICANT: Curtiss, Linda K.

TITLE OF INVENTION: Lipoprotein Assays Using Antibodies to a

TITLE OF INVENTION: Pan Native Epitope and Recombinant Antigens

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESS:

ADDRESSEE: Dressler, Goldsmith, Shore, Sutker &

ADDRESSEE: Milnamow, Ltd.

STREET: 180 No. 5408038th Stetson, Suite 4700

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/959,946

FILING DATE: 19921008

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/901,706

FILING DATE: 18-JUN-1992

ATTORNEY/AGENT INFORMATION:

NAME: Gamson, Edward P.

REGISTRATION NUMBER: 29,381

REFERENCE/DOCKET NUMBER:

TELECOMMUNICATION INFORMATION:

TELEPHONE: (312)616-5400

TELEFAX: (312)616-5460

; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-959-946-5

Query Match 10.3%; Score 91; DB 1; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.047;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 96 KQVPYLDLDFQKKWQEE 111
Db 1 KQVPYLDLDFQKKWQEE 16

RESULT 12
US-08-333-577-5
; Sequence 5, Application US/08333577
; Patent No. 5786206
; GENERAL INFORMATION:
; APPLICANT: Smith, Richard K.
; APPLICANT: Koduri, Raju
; APPLICANT: Young, Stephen G.
; APPLICANT: Witztum, Joseph L.
; APPLICANT: Curtiss, Linda K.
; TITLE OF INVENTION: Lipoprotein Assays Using Antibodies to a
; TITLE OF INVENTION: Pan Native Epitope and Recombinant Antigens
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dressler, Goldsmith, Shore, Sutker &
; ADDRESSEE: Milnamow, Ltd.
; STREET: 180 No. 5786206th Stetson, Suite 4700
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/333,577
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Gamson, Edward P.
; REGISTRATION NUMBER: 29,381
; REFERENCE/DOCKET NUMBER: SCRF 234.0
; TELEPHONE: (312)616-5400
; TELEFAX: (312)616-5460
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-333-577-5

Query Match 10.3%; Score 91; DB 1; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.047;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 96 KQVPYLDLDFQKKWQEE 111
Db 1 KQVPYLDLDFQKKWQEE 16

RESULT 13
PCT-US92-08634-5

; Sequence 5, Application PC/TUS9208634
; GENERAL INFORMATION:
; APPLICANT: Smith, Richard K.
; APPLICANT: Koduri, Raju
; APPLICANT: Young, Stephen G.
; APPLICANT: Witztum, Joseph L.
; APPLICANT: Curtiss, Linda K.
; TITLE OF INVENTION: Lipoprotein Assays Using Antibodies to a
; TITLE OF INVENTION: Pan Native Epitope and Recombinant Antigens
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dressler, Goldsmith, Shore, Sutker &
; ADDRESSEE: Milnamow, Ltd.
; STREET: 180 North Stetson, Suite 4700
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60601

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/08634
; FILING DATE: 19921009
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/901,706
; FILING DATE: 18-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Gamson, Edward P.
; REGISTRATION NUMBER: 29,381
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312)616-5400
; TELEFAX: (312)616-5460

; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US92-08634-5

Query Match 10.3%; Score 91; DB 5; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.047;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 96 KQVPYLDLDFQKKWQEE 111
Db 1 KQVPYLDLDFQKKWQEE 16

RESULT 14
US-09-079-030-129
; Sequence 129, Application US/09079030
; Patent No. 6635623
; GENERAL INFORMATION:
; APPLICANT: Guevera, Jr., Juan G.
; APPLICANT: Hoogveen, Ron C.
; APPLICANT: Moore, Paul J.
; TITLE OF INVENTION: LIPOPROTEINS AS NUCLEIC ACID DELIVERY
; TITLE OF INVENTION: VECTORS FOR TRANSFECTION OF EUKARYOTIC CELLS
; NUMBER OF SEQUENCES: 229
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:

Search completed: December 21, 2004, 13:52:13
Job time : 40 secs

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/079,030
; FILING DATE: Concurrently Herewith
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: McMillian, Nabeela R.
; REGISTRATION NUMBER: P-43,363
; REFERENCE/DOCKET NUMBER: ARAG:003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 129:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 114 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
;
US-09-079-030-129

Query Match 10.1%; Score 89; DB 4; Length 114;
Best Local Similarity 25.0%; Pred. No. 0.82;
Matches 26; Conservative 24; Mismatches 28; Indels 26; Gaps 3;

QY 75 LEKETGLQEMSKDLEEVKAKVQPYLDDPQKKQWEMELYRQKVEPLRAELQEGARQKL 134
Db 3 LGQSTBELRVRLASHLURKURKLRDADDLQKR-----LAVY-----QAGAREGAERGL 51

QY 135 HELQEKLSPL-----GEEMDRARAHVDALRTHL 163
Db 52 SAIRERLGPLVEQGRVRAATVGSLAGOPLQERAAQANGERLARFM 95

RESULT 15
US-09-583-110-3496
; Sequence 3496, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 3496
; LENGTH: 129
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
;
US-09-583-110-3496

Query Match 9.9%; Score 88; DB 4; Length 129;
Best Local Similarity 25.8%; Pred. No. 1.2;
Matches 25; Conservative 27; Mismatches 41; Indels 4; Gaps 1;

QY 62 EQLGPTVTFQFWMLEKETGLQEMSKDLEEVKAKVQPYLDDPQKKQWEMELYRQKVEP 121
Db 27 KQVCSQAQDFLDDREDPEYAKEQVCVKLTVEVKEQA---TDFVLTKEQVESGEITVDS 82

QY 122 LRAELQEGARQKLHELQEKLSPLGEEMDRARAHVDA 158
Db 83 ILAQAKSYAFQATEASKQNQLNKLQEQWKEAEALDDSD 119
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2004, 13:19:15 ; Search time 38 Seconds
(without alignments)
303.842 Million cell updates/sec

Title: US-09-803-918A-2_COPY_25_144

Perfect score: 631
Sequence: 1 DEPPQSPWDRVKDLATVYVD.....LDDFKKQWEMELYRQKVE 120

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 52135

Minimum DB seq length: 0
Maximum DB seq length: 120

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	106	16.8	34	2 S67972	apolipoprotein AI
2	106	16.8	36	2 A56866	apolipoprotein A-I
3	101	16.0	20	2 A05313	apolipoprotein A-I
4	81	12.8	118	2 B69818	conserved hypothetical
5	76.5	12.1	79	2 G86722	hypothetical prote
6	76	12.0	99	2 B46598	ski-related protei
7	69	10.9	79	2 S53443	kinetoplastid memb
8	66	10.5	113	2 C89923	hypothetical prote
9	65.5	10.4	92	2 S53442	kinetoplastid memb
10	65	10.3	113	2 I40399	flagellar protein
11	64.5	10.2	104	2 A60094	neurofilament-M ho
12	64.5	10.2	115	2 C86901	hypothetical prote
13	63.5	10.1	109	2 F71511	hypothetical prote
14	63	10.0	107	1 H64112	virulence-associat
15	62	9.8	106	2 C90261	hypothetical prote
16	61.5	9.7	96	2 I68742	integral membrane
17	61.5	9.7	119	2 B82304	hypothetical prote
18	61	9.7	75	2 C90325	hypothetical prote
19	61	9.7	75	2 C90342	hypothetical prote
20	61	9.7	108	2 A59010	antifreeze protein
21	60.5	9.6	73	2 H81298	probable molybdopt
22	60.5	9.6	107	2 D95153	hypothetical prote
23	60	9.5	113	2 H97483	hypothetical prote
24	60	9.5	119	2 D75111	hypothetical prote
25	60	9.5	119	2 E70049	hypothetical prote
26	59.5	9.4	76	2 D82445	hypothetical prote
27	59.5	9.4	99	2 JU0038	nonhistone chromos
28	59.5	9.4	109	2 F81669	conserved hypothet
29	59.5	9.4	117	2 H71978	hypothetical prote

30	59.5	9.4	120	2 AB1425	hypothetical prote
31	59	9.4	79	2 AE0012	conserved hypothet
32	59	9.4	102	2 D97844	hypothetical prote
33	58.5	9.3	76	2 JC5734	apolipoprotein A-I
34	58.5	9.3	79	2 AB3643	hypothetical cytos
35	58.5	9.3	85	2 C65120	hypothetical 10.0
36	58.5	9.3	93	2 AF3532	transposase BME110
37	58.5	9.3	93	2 AI3537	hypothetical prote
38	58.5	9.3	111	2 B90187	ribosomal protein
39	58.5	9.3	120	2 AD1468	conserved hypothet
40	58	9.2	71	2 A69154	hypothetical prote
41	58	9.2	88	2 A72222	hypothetical prote
42	58	9.2	103	2 B69466	hypothetical prote
43	58	9.2	105	2 C83942	hypothetical prote
44	58	9.2	109	2 E34510	homeotic protein E
45	58	9.2	110	2 E70157	conserved hypothet

ALIGNMENTS

RESULT 1

S67972
apolipoprotein AI - goose (fragment)
C:Species: Anser anser (domestic goose)
C:Date: 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C:Accession: S67972
R:Hernier, D.; Sellier, N.; Rousselot-Pailley, D.; Forgez, P.
Eur. J. Biochem. 234, 586-591, 1995
A:Title: Characterization of apolipoproteins B-100, AI and C from plasma lipoprotein in
A:Reference number: S67972; MUID:96128192; PMID:8536707
A:Accession: S67972
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-34 <HER>
A:Cross-references: UNIPROT:Q9PRR6
C:Superfamily: apolipoprotein A-I

Query Match 16.8%; Score 106; DB 2; Length 34;
Best Local Similarity 54.5%; Pred. No. 0.03;
Matches 18; Conservative 9; Mismatches 6; Indels 0; Gaps 0;

QY 2 EPQQSPWDRVKDLATVYVDLKDSDGRDYSQFE 34
: ||| ||| ||| ||| : ||| ||| |||
Db 1 DEFPALDLRLKDLVDVLTETVKASGKDYLAQFE 33

RESULT 2

A56866
apolipoprotein A-I - Japanese quail (fragment)
C:Species: Coturnix coturnix japonica (Japanese quail)
C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 09-Jul-2004
C:Accession: A56866
R;Oku, H.; Ishikawa, M.; Nagata, J.; Toda, T.; Chinen, I.
Biochim. Biophys. Acta 1167, 22-28, 1993
A:Title: Lipoprotein and apoprotein profile of Japanese quail.
A:Reference number: A56866; MUID:93213845; PMID:8461329
A:Accession: A56866
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-36 <OKU>
A:Cross-references: UNIPROT:P32918
A:Note: sequence extracted from NCBI backbone (NCBI:P128831)
A:Note: this protein was found primarily as a 26K apoprotein
C:Superfamily: apolipoprotein A-I

Query Match 16.8%; Score 106; DB 2; Length 36;
Best Local Similarity 50.0%; Pred. No. 0.032;
Matches 18; Conservative 10; Mismatches 8; Indels 0; Gaps 0;

QY 2 EPQQSPWDRVKDLATVYVDLKDSDGRDYSQFE 37
: ||| ||| ||| ||| : ||| ||| |||
Db 1 DDPQTPDLDRIMLDVLTETVKASGKDAISQFE 36

```
RESULT 3
A05313
apolipoprotein A-I - red guenon (fragment)
N/Alternate names: apo-A-I
C:Species: Erythrocybus patas (red guenon, husear)
C>Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 09-Jul-2004
C/Accession: A05313
R:Mahey, R.W.; Weisgraber, K.H.; Innerarity, T.; Brewer Jr., H.B.
Biochemistry 15, 1928-1933, 1976
A/Title: Characterization of the plasma lipoproteins and apoproteins of the Erythrocybus
A/Reference number: A90395; MUID:76184721; PMID:178359
A/Accession: A05313
A/Molecule type: protein
A/Residues: 1-20 <MAH>
A/Cross-references: UNIPROT:P18647
C:Superfamily: apolipoprotein A-I
C:Keywords: cholesterol metabolism; HDL; intestine; lipid binding; lipid transport; lipid

Query Match 16.0%; Score 101; DB 2; Length 20;
Best Local Similarity 85.0%; Pred. No. 0.04;
Matches 17; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 DEPPQSPWDRVKDLATVYVD 20
Db 1 DEPPQTPWDRVKDLATVYVE 20

RESULT 4
B69818
conserved hypothetical protein yhaH - Bacillus subtilis
C:Species: Bacillus subtilis
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C/Accession: B69818
R:Kunst, P.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte
C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuani, V.; Carter, N.M.; Ch
A.; Ehrlich, S.D.; Emmsrson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A/Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holtsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A/Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Maguda, S.; Maueel
Y.; M.; Ogawa, K.; Ogiwara, C.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon,
A/Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron
akeuchi, M.; Takakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A/Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A/Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A/Reference number: A69580; MUID:98044033; PMID:9384377
A/Accession: B69818
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-118 <KUN>
A/Cross-references: UNIPROT:O07516; GB:299109; GB:AL009126; NID:G2633260; PIDN:CAB12840.
A/Experimental source: strain 168
C:Genetics:
A/Gene: yhaH

Query Match 12.8%; Score 81; DB 2; Length 118;
Best Local Similarity 26.2%; Pred. No. 11;
Matches 22; Conservative 22; Mismatches 24; Indels 16; Gaps 3;

QY 39 GQNLKLLDNWDSVTSTFSKLRQLGPGVTQEFWDNLEKTEGLR--QEMSKDLBEVKAK 96
Db 30 GQRLKEMKNTYDSPEETIKRLKS-----DGLALKDQLIKAAKSTVDKVGGE 79

QY 97 VQPYLDDFKQWQEMELVRQKVE 120
Db 80 LQTSI----KRWQBEIKPHQDDIQ 99
```

```
RESULT 5
G86722
hypothetical protein yhjA [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C:Species: Lactococcus lactis subsp. lactis
C>Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004
C/Accession: G86722
R:Botolin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Ehrli
Genome Res. 11, 731-753, 2001
A/Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s
A/Reference number: A86625; MUID:21235186; PMID:11337471
A/Accession: G86722
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-79 <STO>
A/Cross-references: UNIPROT:Q9CHE9; GB:AE005176; PID:gi2723703; PIDN:AAK04891.1; GSPDB:
A/Experimental source: strain IL1403
C:Genetics:
A/Gene: yhjA

Query Match 12.1%; Score 76.5; DB 2; Length 79;
Best Local Similarity 29.1%; Pred. No. 15;
Matches 23; Conservative 16; Mismatches 37; Indels 3; Gaps 1;

QY 42 LNLKLLDNWDSVTSTFSKLRQLGPGVTQEFWDNLEKTEGLRQEMSKDLBEVKAKVQPYL 101
Db 3 LNDKLLDATKDKVS---GKVKETTKVTGDEKLEAKGKTEGLMGKAKEGLENIKDKASDLA 59

QY 102 DDFQKKQWQEMELVRQKVE 120
Db 60 EDVAEKFNFTVDSVKGKNE 78

RESULT 6
B46598
ski-related protein SnO - chicken (fragments)
C:Species: Gallus gallus (chicken)
C>Date: 03-Mar-1994 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C/Accession: B46598
R:Nagase, T.; Nomura, N.; Ishii, S.
J. Biol. Chem. 268, 13710-13716, 1993
A/Title: Complex formation between proteins encoded by the ski gene family.
A/Reference number: A46598; MUID:93293901; PMID:8514802
A/Accession: B46598
A/Status: preliminary; not compared with conceptual translation
A/Molecule type: mRNA
A/Residues: 1-56; 57-99 <NAG>
A/Cross-references: UNIPROT:Q9PS83
A/Note: sequence extracted from NCBI backbone (NCBIP:134568, NCBIP:134572)
C:Superfamily: ski transforming protein
C:Keywords: DNA binding; tandem repeat

Query Match 12.0%; Score 76; DB 2; Length 99;
Best Local Similarity 31.0%; Pred. No. 22;
Matches 27; Conservative 20; Mismatches 20; Indels 20; Gaps 5;

QY 41 QNLKLLDNWDSVTSTFSKLRQLGPGVTQEFWDNLEKTEGLRQEMSKDLBEVKAKVQPY 100
Db 17 QMEVKMLSS-----SKMKEL---TEE-QQNLQKLESLQNEHAQRMEEQLAEQR 64

QY 101 LD-----DFQKKWQEMELVRQKVE 120
Db 65 LDHAEDRQELQDELQRQREA-RQKLE 90

RESULT 7
S53443
kinetoplastid membrane protein-11 - Leishmania donovani
C:Species: Leishmania donovani
C>Date: 01-Aug-1995 #sequence_revision 01-Sep-1995 #text_change 07-May-1999
C/Accession: S53443
R:Jardim, A.; Funk, V.; Caprioli, R.M.; Olafson, R.W.
Biochem. J. 305, 307-313, 1995
A/Title: Isolation and structural characterization of the Leishmania donovani kinetoplas
```

A;Reference number: S53443; MUID:95126922; PMID:7826346
A;Accession: S53443
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-79 <JAR>

Query Match 10.9%; Score 69; DB 2; Length 79;
Best Local Similarity 29.0%; Pred. No. 59;
Matches 20; Conservative 15; Mismatches 26; Indels 8; Gaps 3;

QY 51 DSVTSTFS-KLRQLGPTQEFWDNLEKTEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQ 109
DB 2 DRLDEEFNRKMQSQ-----NEFADKPDEST--LSPENRHYEKFERMIXEHTKFNKQVH 54

QY 110 EEMELYRQK 118
DB 55 EHESEFKQK 63

RESULT 8
C89923
hypothetical protein [imported] - Staphylococcus aureus (strain N315)
C;Species: Staphylococcus aureus
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C;Accession: C89923
R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogura, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A;Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A;Reference number: A89758; MUID:21311952; PMID:11418146
A;Accession: C89923
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-113 <KUR>
A;Cross-references: UNIPROT:Q99U38; GB:BA000018; PID:g13701249; PIDN:BAB42544.1; GSPDB:G
A;Experimental source: strain N315
C;Genetics:
A;Gene: SA1284

Query Match 10.5%; Score 66; DB 2; Length 113;
Best Local Similarity 34.1%; Pred. No. 1.5e+02;
Matches 14; Conservative 9; Mismatches 14; Indels 4; Gaps 1;

QY 67 VTQEFWDNLEKTEGLRQEMSKDLEEVKAKVQPYLDDFQKK 107
DB 4 ITKEVDNLEQEDVFAKNAKLGSSSEAK----PYLDEYHSK 40

RESULT 9
S53442
kinetoplastid membrane protein-11 - Leishmania donovani
C;Species: Leishmania donovani
C;Date: 01-Aug-1995 #sequence_revision 01-Sep-1995 #text_change 07-May-1999
C;Accession: S53442
R;Jardim, A.; Hanson, S.; Ullman, B.; McCubbin, W.D.; Kay, C.M.; Olafson, R.W.
Biochem. J. 305, 315-320, 1995
A;Title: Cloning and structure-function analysis of the Leishmania donovani kinetoplasti

A;Reference number: S53442; MUID:95126923; PMID:7826347
A;Accession: S53442
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-92 <JAR>

Query Match 10.4%; Score 65.5; DB 2; Length 92;
Best Local Similarity 24.3%; Pred. No. 1.3e+02;
Matches 18; Conservative 17; Mismatches 32; Indels 7; Gaps 2;

QY 50 WDSVTSTFSKLRQLGPTQEFWDNLEKTEGLRQEMSKDLEEVKAKVQPYLDDF 104
DB 5 YEEFSKLRDLDEEFNRKMQSQNAKPFADKPDEST--LSPENRHYEKFERMIXEHTKFP 62

QY 105 QKKWQEMELYRQK 118

DB 63 NKKMHEHSEHFKQK 76

RESULT 10
I40399
flagellar protein flit - Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 09-Jul-2004
C;Accession: I40399; P69625
R;Chen, L.; Helmann, J.D.
J. Bacteriol. 176, 3093-3101, 1994
A;Title: The Bacillus subtilis sigma D-dependent operon encoding the flagellar proteins
A;Reference number: I40396; MUID:94252974; PMID:8195064
A;Accession: I40399
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-113 <RES>
A;Cross-references: UNIPROT:P39740; EMBL:Z31376; NID:ig499379; PIDN:CAA83250.1; PID:ig499
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berthe
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Puma, S.; Galizzi, A.; Gall
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mauee
Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetcel
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Sero
akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, K.; Tosato, V.; Uchiyama
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Accession: F69625
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-113 <KUN>
A;Cross-references: GB:Z99122; GB:AL009126; NID:g2636029; PIDN:CAB15549.1; PID:el184438
A;Experimental source: strain 168
C;Genetics:
A;Gene: flit

Query Match 10.3%; Score 65; DB 2; Length 113;
Best Local Similarity 24.0%; Pred. No. 1.8e+02;
Matches 24; Conservative 26; Mismatches 28; Indels 22; Gaps 6;

QY 19 VDVLKDSGRDYVSQFEGSALGKQLNLKLDNWDVSTFTSKLREQLG-PVTQF----- 70
DB 4 IDQLVTETKSLMSLHIQNTPESEDL-LKQIEDF---VATRSLEIQEISLPSEERKQMKL 59

QY 71 --FWNVL-EKTEGLRQEMSKDLEEVKAK-----VQPY 100
DB 60 ILTWDLQIVKEMERLKSQSIATELQOMKRVNHTTYLNPY 99

RESULT 11
A60094
neurofilament-M homolog XIF6 - African clawed frog (fragment)
C;Species: Xenopus laevis (African clawed frog)
C;Date: 03-Mar-1993 #sequence_revision 03-Mar-1993 #text_change 09-Jul-2004
C;Accession: A60094
R;Sharpe, C.R.
Development 103, 269-277, 1988
A;Title: Developmental expression of a neurofilament-M and two vimentin-like genes in x
A;Reference number: A60094; MUID:89136789; PMID:3224553
A;Accession: A60094
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-104 <SHA>
A;Cross-references: UNIPROT:O13099; UNIPROT:O13098
C;Superfamily: cytoskeletal keratin

Query Match 10.2%; Score 64.5; DB 2; Length 104;
Best Local Similarity 24.0%; Pred. No. 1.8e+02;
Matches 24; Conservative 18; Mismatches 35; Indels 23; Gaps 4;

QY 25 SGRDYVSQFEGSALGKQLNKLDDNWDVSTSTFSKLRQLGVPVT-----QEFWDN 74
DB 15 SAKDEIAEYR-----RQIHSKTVE-LESVRGKTESLERQLNDIERHGHDLTSYQETQQ 68

QY 75 LEKETEGLRQEMSKOLEEVKAKVQPYLDDFQKKWQEMEL 114
DB 69 LDNELRGTKWENSRHLRE-----YQDLLNVKQALDIEI 101

RESULT 12
C86901
hypothetical protein ywjb [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C:Species: Lactococcus lactis subsp. lactis
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004
C:Accession: C86901
R:Boletín, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Ehrlich
Genome Res. 11, 731-753, 2001
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis se
A:Reference number: A86625; MUID:21235186; PMID:11337471
A:Accession: C86901
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-115 <STO>
A:Cross-references: UNIPROT:O9CDK8; GB:AE0051176; PID:g12725277; PIDN:AAK06309.1; GSPDB:G
A:Experimental source: strain IL1403
C:Genetics:
A:Gene: ywjb

Query Match 10.2%; Score 64.5; DB 2; Length 115;
Best Local Similarity 20.2%; Pred. No. 2e+02;
Matches 21; Conservative 24; Mismatches 44; Indels 15; Gaps 3;

QY 13 DLATVYVDVLKDSGRDYVSQFEGSALGKQLNKLDDNWDVSTSTFSKLRQLGVPVTOE-- 70
DB 5 DLVLDLVEKIK--SLDYVKDFQQAETALMANQELF-----KQAEEMKALQKEAV 51

QY 71 FWDNLEKETEGRLRQEMSKOLEEVKAKVQPYLDDFQKKWQEMEL 114
DB 52 LVQKIDMQAYKTSQSAQVIEKIKLHPLVEDYATKLEDVNDL 95

RESULT 13
F71511
hypothetical protein CT466 - Chlamydia trachomatis (serotype D, strain UW3/Cx)
C:Species: Chlamydia trachomatis
C:Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 09-Jul-2004
C:Accession: F71511
R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell,
Science 282, 754-759, 1998
A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia tra
A:Reference number: A71570; MUID:99000809; PMID:9784136
A:Accession: F71511
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-109 <ARN>
A:Cross-references: UNIPROT:O84472; GB:AE001320; GB:AE001273; NID:g3328891; PIDN:AAC6806
A:Experimental source: serotype D, strain UW-3/Cx
C:Genetics:
A:Gene: CT466
C:Superfamily: conserved hypothetical protein CP0165

Query Match 10.1%; Score 63.5; DB 2; Length 109;
Best Local Similarity 24.7%; Pred. No. 2.3e+02;
Matches 22; Conservative 25; Mismatches 31; Indels 11; Gaps 4;

QY 30 VSQFEGSALGKQLNKLDDNWDVSTSTFSKLRQLGVPVTOEFWDNLEKETEGRLRQEMSKD 89
DB 23 LTQVEASTSPEDL-IKVLROKKTLLSCIEKVDHQI-----KKFRDSF---SLALPQEQEE 74

QY 90 LEEVKAKVQPYLDDFQKKW---QSEMELY 115
DB 75 LEEIRSVIQRILETDKKNYCIRKRELRTY 103

RESULT 14
H64112
virulence-associated protein vapA homolog HI1251 - Haemophilus influenzae (strain Rd KW2
C:Species: Haemophilus influenzae
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: H64112
R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.
; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghegan, N.S.M.
Science 269, 496-512, 1995
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A:Reference number: A64000; MUID:95350630; PMID:7542800
A:Accession: H64112
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-107 <TIGR>
A:Cross-references: UNIPROT:Q57089; GB:U32805; GB:L42023; NID:g1574180; PIDN:AAC22901.1;
C:Superfamily: virulence-associated protein vapA
C:Keywords: DNA binding; transcription regulation

Query Match 10.0%; Score 63; DB 1; Length 107;
Best Local Similarity 23.9%; Pred. No. 2.4e+02;
Matches 22; Conservative 21; Mismatches 31; Indels 18; Gaps 3;

QY 10 RVKDLATVYVDVLKDSGRDYVSQFEGSALGKQLNKLDDNWDVSTSTFSKLRQLGVPVTO 69
DB 24 KISDLAQI-LDVHRTASNIYN--NSSRITLLEMAVKLAKVFDT-----TP 65

QY 70 EFDNLEKETEGRLRQEMSKOLEEVKAKVQPYL 101
DB 66 EFWNLQTRIDLWDLHNKRFQOSLANVKPAL 97

RESULT 15
C90261
hypothetical protein SSO1086 [imported] - Sulfolobus solfataricus
C:Species: Sulfolobus solfataricus
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004
C:Accession: C90261
R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-
arrett, R.A.; Ragan, M.A.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.
submitted to GenBank, April 2001
A:Description: Sulfolobus solfataricus complete genome.
A:Reference number: A99139
A:Accession: C90261
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-106 <KUP>
A:Cross-references: UNIPROT:Q97Z50; GB:AE006641; NID:g13814276; PIDN:AAK41346.1; GSPDB:G
C:Genetics:
A:Gene: SSO1086

Query Match 9.8%; Score 62; DB 2; Length 106;
Best Local Similarity 32.5%; Pred. No. 2.9e+02;
Matches 25; Conservative 15; Mismatches 21; Indels 16; Gaps 5;

QY 48 DNWDVST-----TFSKLRQLGVPVTOEFWDNLEKETEGRLRQEMSKOLEEVKAKVOP 99
DB 35 NNWVSVSAERFRSNDPSINKVKAQVATHE--EDLKAIEELRSSGIK-IEEVKE--IP 89

QY 100 YLDDF---QKKWQEME 113
DB 90 FPRKPIEGRKXIOBEFD 106

Search completed: December 21, 2004, 13:31:26
Job time : 40 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2004, 13:19:55 ; Search time 190 Seconds
(without alignments)
363.395 Million cell updates/sec

Title: US-09-803-918A-2_COPY_25_144

Perfect score: 631
Sequence: 1 DBPQSPWDRKDLATVYVD.....LDDFKKWOEMELRYQKVE 120

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 393250

Minimum DB seq length: 0
Maximum DB seq length: 120

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot 02:.*
1: uniprot_sprot:.*
2: uniprot_trembl:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	247	39.1	56	2	O02762 oviv aries
2	226	35.8	79	2	O6LD50 mus sp. apo
3	226	35.8	79	2	AAB35539 mus sp. a
4	217	34.4	82	2	O29248 sus scrofa
5	141	22.3	26	2	O9UCT8
6	106	16.8	34	2	O9PRR6
7	101	16.0	20	1	APAL_ERYPA
8	101	16.0	30	2	O9QV04
9	85	13.5	107	1	APE_MACMU
10	83	13.2	107	1	APE_SALSC
11	81	12.8	118	2	O07516
12	77.5	12.3	96	2	O7P280
13	76.5	12.1	79	2	O9CHE9
14	76	12.0	99	2	O9PSS3
15	76	12.0	118	2	O68959
16	75	11.9	117	2	O68958
17	74.5	11.8	102	2	O9PXP4
18	73	11.6	76	2	O6WHJ6
19	73	11.6	76	2	AAQ64377
20	73	11.6	93	2	O8IR54
21	73	11.6	93	2	AAT31321
22	73	11.6	113	2	O8IIF8
23	72	11.4	92	2	O7YZC6
24	72	11.4	102	2	O29258
25	70.5	11.2	92	2	O9GU60
26	70.5	11.2	114	2	O9ZFE1
27	70	11.1	92	1	KM11_TRYBB
28	70	11.1	92	2	O7YZC5
29	69.5	11.0	92	2	O9XZHS
30	69	10.9	104	2	O54892
31	68	10.8	90	2	O9LCR0

32	68	10.8	101	2	Q6BMU6	Q6bmuf debaryomyce
33	68	10.8	105	2	O85812	O85812 streptococc
34	67.5	10.7	92	2	O43974	O43974 leishmania
35	67	10.6	78	2	O9BHF1	O9bhf1 streptococc
36	67	10.6	92	2	O8IS88	O8is88 trypanosoma
37	67	10.6	95	2	P89597	P89597 human immun
38	66.5	10.5	92	1	K11B_LEIIN	K11b_leishmania
39	66.5	10.5	92	1	K11C_LEIIN	K25297 leishmania
40	66.5	10.5	92	1	KM11_LEITR	O21436 leishmania
41	66.5	10.5	92	2	O6RV15	O6rv15 leishmania
42	66.5	10.5	92	2	AAR84616	Aar84616 leishmani
43	66	10.5	80	2	Q6NK47	Q6nk47 corynebacte
44	66	10.5	80	2	CAE48695	CaE48695 corynebac
45	66	10.5	92	1	KM11_LEIDO	Q36736 leishmania

ALIGNMENTS

```
RESULT 1
O02762 PRELIMINARY; PRT; 56 AA.
AC O02762;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Apolipoprotein A1 (Fragment).
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99098454; PubMed=9883985;
RA Robertson J.A., Bhattacharyya S., Ing N.H.;
RT "tamoxifen up-regulates oestrogen receptor-alpha, c-fos and
RT glyceraldehyde 3-phosphate-dehydrogenase mRNAs in ovine endometrium.";
RL J. Steroid Biochem. Mol. Biol. 67:285-292 (1998).
DR EMBL; U94720; AAB57840.1; -
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0008289; F:lipid binding; IEA.
DR GO; GO:0006869; P:lipid transport; IEA.
DR GO; GO:0042157; P:lipoprotein metabolism; IEA.
DR InterPro; IPR000074; Apolipoprotein.
DR InterPro; IPR009074; Apolipo_A_E_C3.
DR Pfam; PF01442; Apolipoprotein; 1.
KW Lipoprotein.
FT NON_TER 1
FT NON_TER 56
SQ SEQUENCE 56 AA; 6617 MW; 2AB38E08F1E8F1BC CRC64;
Query Match 39.1%; Score 247; DB 2; Length 56;
Best Local Similarity 82.1%; Pred. No. 1.4e-12;
Matches 46; Conservative 3; Mismatches 7; Indels 0; Gaps 0;
```

```
QY 47 LDNWDSVTSFKSLRQLGPVTQEFWDNLEKETEGLRQBSKDLBVKAKVQPYLD 102
Db 1 LDNWDSLASTLSKVRQLGPVTQEFWDNLEKETASURQEMHKDLBARKVQPYLD 56
```

RESULT 2

```
O6LD50 PRELIMINARY; PRT; 79 AA.
ID Q6LD50
AC Q6LD50;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Apolipoprotein A1 homolog protein (Fragment).
GN Name=apolipoprotein A1 homolog;
OS Mus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
```

```

OX NCBI_TaxID=10095;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96062440; PubMed=7488287;
RA Merrill J.T., Rivkin E., Shen C., Lahita R.G.;
RT "Selection of a gene for apolipoprotein A1 using autoantibodies from a
RT patient with systemic lupus erythematosus.";
RL Arthritis Rheum. 38:1655-1659(1995).
DR EMBL; S80442; AAB35539.1; -.
DR InterPro; IPR009074; Apolipo_A_E_C3.
KW Lipoprotein.
FT NON_TER
SQ SEQUENCE 79 AA; 9583 MW; 33CA72DA854A150A CRC64;

Query Match 35.8%; Score 226; DB 2; Length 79;
Best Local Similarity 80.0%; Pred. No. 1e-10;
Matches 40; Conservative 8; Mismatches 2; Indels 0; Gaps 0;

QY 70 EFWDNLEKETEGRLQEMSKDLSEVKAKVQPYLDDFQKKQWEMELYRQKV 119
Db 1 DFWDNLEKETDVRQEMNKDLSEVKQKQVPLYLDEFQKKWKEDVELYRQKV 50

RESULT 3
AAB35539
ID AAB35539 PRELIMINARY; PRT; 79 AA.
AC AAB35539
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE Apolipoprotein A1 homolog (Fragment).
OS Mus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10095;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96062440; PubMed=7488287;
RA Merrill J.T., Rivkin E., Shen C., Lahita R.G.;
RT "Selection of a gene for apolipoprotein A1 using autoantibodies from a
RT patient with systemic lupus erythematosus.";
RL Arthritis Rheum. 38:1655-1659(1995).
DR EMBL; S80442; AAB35539.1; -.
FT NON_TER
SQ SEQUENCE 79 AA; 9583 MW; 33CA72DA854A150A CRC64;

Query Match 35.8%; Score 226; DB 2; Length 79;
Best Local Similarity 80.0%; Pred. No. 1e-10;
Matches 40; Conservative 8; Mismatches 2; Indels 0; Gaps 0;

QY 70 EFWDNLEKETEGRLQEMSKDLSEVKAKVQPYLDDFQKKQWEMELYRQKV 119
Db 1 DFWDNLEKETDVRQEMNKDLSEVKQKQVPLYLDEFQKKWKEDVELYRQKV 50

RESULT 4
Q29248
ID Q29248 PRELIMINARY; PRT; 82 AA.
AC Q29248
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Apolipoprotein A-I (Fragment).
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Small intestine;
RX MEDLINE=96327607; PubMed=8672129;
RA Winteroe A.K., Fredholm M., Davies W.;
RT "Evaluation and characterization of a porcine small intestine cDNA
```

```

RT library.";
RL Mamm. Genome 7:509-517(1996).
DR EMBL; F14858; CAA23298.1; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0008289; F:lipid binding; IEA.
DR GO; GO:0006869; P:lipid transport; IEA.
DR GO; GO:0042157; P:lipoprotein metabolism; IEA.
DR InterPro; IPR000074; Apolipoprotein.
DR InterPro; IPR009074; Apolipo_A_E_C3.
DR Pfam; PF01442; Apolipoprotein; 1.
KW Lipoprotein.
FT NON_TER
SQ SEQUENCE 82 AA; 9168 MW; 24625C65CBFFEDD8 CRC64;

Query Match 34.4%; Score 217; DB 2; Length 82;
Best Local Similarity 70.7%; Pred. No. 5.5e-10;
Matches 41; Conservative 5; Mismatches 12; Indels 0; Gaps 0;

QY 2 EPPOSQPDWRVLDLATVYVDVLKSGRDYVVSQFEGSALGKQLNLKLLDNWDSVTSTFSK 59
Db 25 DDPXSPDWRVLDLATVYVDVLKSGRDYVVSQFEGSALGKHLNLKLLXNDSLGXTFTK 82

RESULT 5
Q5UCT8
ID Q5UCT8 PRELIMINARY; PRT; 26 AA.
AC Q5UCT8
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Sperm activating protein subunit I, apolipoprotein A1, SPAP subunit I
DE (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=91369902; PubMed=1909888;
RA Akerlof E., Jornvall H., Slotte H., Pousette A.;
RT "Identification of apolipoprotein A1 and immunoglobulin as components
RT of a serum complex that mediates activation of human sperm motility.";
RL Biochemistry 30:8986-8990(1991).
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0008289; F:lipid binding; IEA.
DR GO; GO:0006869; P:lipid transport; IEA.
DR GO; GO:0042157; P:lipoprotein metabolism; IEA.
DR InterPro; IPR000074; Apolipo_A_E_C3.
DR InterPro; IPR009074; Apolipo_A_E_C3.
DR Pfam; PF01442; Apolipoprotein; 1.
FT NON_TER
SQ SEQUENCE 26 AA; 2930 MW; 64921A333E768D27 CRC64;

Query Match 22.3%; Score 141; DB 2; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.00017;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DEPPQSPDWRVLDLATVYVDVLKDSG 26
Db 1 DEPPQSPDWRVLDLATVYVDVLKDSG 26

RESULT 6
Q9PRR6
ID Q9PRR6 PRELIMINARY; PRT; 34 AA.
AC Q9PRR6
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Apolipoprotein A1 (Fragment).
OS Anser anser (domestic goose).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
OC Archosauria; Aves; Neognathae; Anseriformes; Anatidae; Anser.
OX NCBI_TaxID=8843;
RN [1]
RP SEQUENCE.
RX MEDLINE=96128192; PubMed=8536707;
RA Hermier D., Sellier N., Roussetot-Pailley D., Forgez P.;
RT "Characterization of apolipoproteins B-100, AI and C from plasma
RT lipoprotein in the goose, Anser anser. Evidence for a genetic
RT polymorphism in ApoC-like apolipoproteins.";
RL Eur. J. Biochem. 234:586-591(1995).
DR PIR; S67972; S67972.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0008289; F:lipid binding; IEA.
DR GO; GO:0006869; P:lipid transport; IEA.
DR GO; GO:0042157; P:lipoprotein metabolism; IEA.
DR InterPro; IPR000074; Apolipoprotein.
DR Pfam; PF01442; Apolipoprotein; 1.
DR SEQUENCE 34 AA; 3838 MW; 4BE7DFA02BFIDE91 CRC64;
SQ
Query Match 16.0%; Score 106; DB 2; Length 34;
Best Local Similarity 54.5%; Pred. No. 0.14;
Matches 18; Conservative 9; Mismatches 6; Indels 0; Gaps 0;
QY 2 EPPQSPWDRVKDLATVYVDLKGSRDYVSQFE 34
: ||| ||| ||| ||| : ||| ||| |||
Db 1 DEPPQAPLRLKDLVDVYLETVTASGKDYLAQFE 33
: ||| ||| ||| ||| : ||| ||| |||
RESULT 7
APAI ERYPA
ID APAI ERYPA STANDARD; PRT; 20 AA.
AC P18647;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Apolipoprotein A-I (Apo-AI) (Fragment).
GN Name=APOAI;
OS Erythrocybus patas (Red quonon) (Cercopithecus patas).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC Cercopithecinae; Erythrocybus.
OX NCBI_TaxID=9538;
RN [1]
RP SEQUENCE.
RX MEDLINE=76184721; PubMed=178359;
RA Mahley R.W., Weisgraber K.H., Innerarity T., Brewer H.B. Jr.;
RT "Characterization of the plasma lipoproteins and apoproteins of the
RT Erythrocybus patas monkey.";
RL Biochemistry 15:1928-1933(1976).
CC -!- FUNCTION: Participates in the reverse transport of cholesterol
CC from tissues to the liver for excretion by promoting cholesterol
CC efflux from tissues and by acting as a cofactor for the lecithin
CC cholesterol acyltransferase (LCAT).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Major protein of plasma HDL, also found in
CC chylomicrons.
CC -!- SIMILARITY: Belongs to the apolipoprotein AI/A4/E family.
DR PIR; A05313; A05313.
DR InterPro; IPR009074; Apolipo_A_E_C3.
KW Cholesterol metabolism; Direct protein sequencing; HDL;
KW Lipid transport; Plasma.
FT NON_TER 20 20
SQ SEQUENCE 20 AA; 2387 MW; 9C970997CFC976A CRC64;
Query Match 16.0%; Score 101; DB 1; Length 20;
Best Local Similarity 85.0%; Pred. No. 0.19;
Matches 17; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 DEPPQSPWDRVKDLATVYVD 20
||| ||| ||| ||| |||
Db 1 DEPPQTFWDRVKDLVTVYVE 20
||| ||| ||| ||| |||
RESULT 8
Q9QV04
ID Q9QV04 PRELIMINARY; PRT; 30 AA.
AC Q9QV04;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Apolipoprotein E (Fragment).
OS Rattus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10118;
RN [1]
RP SEQUENCE.
RX MEDLINE=95355312; PubMed=7629028;
RA Motojima K., Goto S.;
RL J. Biochem. 117:597-602(1995).
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0008289; F:lipid binding; IEA.
DR GO; GO:0006869; P:lipid transport; IEA.
DR GO; GO:0042157; P:lipoprotein metabolism; IEA.
DR InterPro; IPR000074; Apolipoprotein.
DR Pfam; PF01442; Apolipoprotein; 1.
DR SEQUENCE 30 AA; 3355 MW; 4A0103DCB6242BF3 CRC64;
SQ
Query Match 16.0%; Score 101; DB 2; Length 30;
Best Local Similarity 67.9%; Pred. No. 0.3;
Matches 19; Conservative 1; Mismatches 8; Indels 0; Gaps 0;
QY 4 PQSPWDRVKDLATVYVDLKGSRDYVS 31
||| ||| ||| ||| ||| : ||| |||
Db 3 PQSQMDXVXDFATVYVDVAVXSGXDYS 30
||| ||| ||| ||| ||| : ||| |||
RESULT 9
APE MACMU
ID APE MACMU STANDARD; PRT; 107 AA.
AC Q28502;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Apolipoprotein E (Apo-E) (Fragment).
GN Name=APOE;
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96225955; PubMed=8635577;
RA Morelli L., Wei L., Amorim A., McDermaid J., Abec C.R., Frangione B.,
RA Walker L.C., Levy E.;
RT "Cerebrovascular amyloidosis in squirrel monkeys and rhesus monkeys:
RT apolipoprotein E genotype.";
RL FEBS Lett. 379:132-134(1996).
CC -!- FUNCTION: Mediates the binding, internalization, and catabolism of
CC lipoprotein particles. It can serve as a ligand for the LDL (apo
CC B/E) receptor and for the specific apo-E receptor (chylomicron
CC remnant) of hepatic tissues.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Secreted in plasma.
CC -!- SIMILARITY: Belongs to the apolipoprotein AI/A4/E family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
```

```
DR EMBL: U52030; AAC50441.1; -.
DR HSSP: P02649; 1NFN.
DR InterPro; IPR000074; Apolipoprotein.
DR Pfam; PF01442; Apolipoprotein; 1.
DR Chylomicron; Glycoprotein; HDL; Heparin-binding; Lipid transport;
KW Plasma; Repeat; VLDL.
FT NON_TER 1
FT DOMAIN 89 99 LDL receptor binding (Potential).
FT DOMAIN 93 96 Heparin-binding (By similarity).
FT REPEAT 11 >107 8 X 22 AA approximate tandem repeats.
FT REPEAT 11 32 1.
FT REPEAT 33 54 2.
FT REPEAT 55 76 3.
FT REPEAT 77 98 4.
FT REPEAT 99 >107 5.
FT NON_TER 107
SQ SEQUENCE 107 AA; 12382 MW; E1D38C32F5AACB23 CRC64;

Query Match 13.5%; Score 85; DB 1; Length 107;
Best Local Similarity 29.0%; Pred. No. 24;
Matches 18; Conservative 15; Mismatches 29; Indels 0; Gaps 0;

QY 58 SKLRQLGPTVQTFWDLNLEKETEGRLQEMSKDLEEVKAKVQPYLDLDFQKKWQEMELYRQ 117
DB 25 SELEQLSPVAETRLARLSKELQAAQARLGADMDVRSVLQYRSEVQMLGQSTEEELRA 84

QY 118 KV 119
DB 85 RL 86

RESULT 10
APE_SAISC STANDARD; PRT; 107 AA.
AC Q28295;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Apolipoprotein E (Apo-E) (Fragment).
GN Name=APOE;
OS Saimiri sciureus (Common squirrel monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Saimiri.
OX NCBI_TaxID=9521;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96225955; PubMed=8635577;
RA Morelli L., Wei L., Amorim A., McDermid J., Abec C.R., Frangione B.,
RA Walker L.C., Levy E.;
RA "Cerebrovascular amyloidosis in squirrel monkeys and rhesus monkeys:
RT apolipoprotein E genotype."
RL FEBS Lett. 379:132-134(1996).
CC -!- FUNCTION: Mediates the binding, internalization, and catabolism of
CC B/E) lipoprotein particles. It can serve as a ligand for the LDL (apo
CC B/E) receptor and for the specific apo-E receptor (chylomicron
CC remnant) of hepatic tissues.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Secreted in plasma.
CC -!- SIMILARITY: Belongs to the apolipoprotein A1/A4/E family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: U52029; AAC50442.1; -.
DR HSSP: P02649; 1NFN.
DR InterPro; IPR000074; Apolipoprotein.
DR Pfam; PF01442; Apolipoprotein; 1.
```

```
DR Pfam: PF01442; Apolipoprotein; 1.
KW Chylomicron; Glycoprotein; HDL; Heparin-binding; Lipid transport;
FT NON_TER 1
FT DOMAIN 89 99 LDL receptor binding (Potential).
FT DOMAIN 93 96 Heparin-binding (By similarity).
FT REPEAT 11 >107 8 X 22 AA approximate tandem repeats.
FT REPEAT 11 32 1.
FT REPEAT 33 54 2.
FT REPEAT 55 76 3.
FT REPEAT 77 98 4.
FT REPEAT 99 >107 5.
FT NON_TER 107
SQ SEQUENCE 107 AA; 12328 MW; FF88CED47BD18P7C CRC64;

Query Match 13.2%; Score 83; DB 1; Length 107;
Best Local Similarity 33.3%; Pred. No. 35;
Matches 16; Conservative 12; Mismatches 20; Indels 0; Gaps 0;

QY 58 SKLRQLGPTVQTFWDLNLEKETEGRLQEMSKDLEEVKAKVQPYLDLDFQ 105
DB 25 SELEQLSPVAETRLARLSKELQAAQARLGADMDVRSRLAQYRSEVQ 72

RESULT 11
O07516 PRELIMINARY; PRT; 118 AA.
ID AC O07516; Q79609;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein yhaH.
GN Name=yhaH; OrderedLocustNames=BSU10000;
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RA Noback M.A., Terpstra P., Holsappel S., Venema G., Bron S.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377; DOI=10.1038/36786;
RA Kunst F., Ogasawara N., Moser I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriess R., Bourrier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Entian K.-D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
RA Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA Hibert H., Holsappel S., Hosono S., Hulio M.F., Itaya M.,
RA Jones L.-M., Joris B., Karamata D., Kashara Y., Klaer-Blanchard M.,
RA Klein C., Kobayashi Y., Koetter P., Koningsstein G., Krogh S.,
RA Kumano M., Kurita K., Lapidus A., Lardinois S., Lauber J.,
RA Lazarevic V., Lee S.M., Levine A., Liu H., Masuda S., Mauel C.,
RA Medigue C., Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S.,
RA Noback M., Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B.,
RA Park S.H., Parro V., Pohl T.M., Portetelle D., Porwollik S.,
RA Prescott A.M., Prescan E., Pujic P., Purnelle B., Rapoport G.,
RA Roy M., Reynolds S., Rieger M., Rivolta C., Rocha E., Roche B.,
RA Rose M., Sadaie Y., Sato T., Scanlan E., Schleich S., Schroeter R.,
RA Scoffone F., Sekiguchi J., Sekowska A., Seter S.J., Serron P.,
RA Shin B.S., Soldo B., Sorokin A., Tacconi E., Takagi T., Takahashi H.,
RA Takemaru K., Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P.,
RA Tognoni A., Tosato V., Uchiyama S., Vandenbol M., Vannier F.,
RA Vassarotti A., Viari A., Wambutt R., Wedler H.,
RA Weitzenegger T., Winters P., Wipat A., Yamamoto H., Yamane K.,
RA Yasumoto K., Yata K., Yoshida K., Yoshikawa H.F., Zumstein E.,
```

```
RA Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RL Nature 390:249-256(1997).
DR EMBL: Y14077; CAA74413.1; -.
DR EMBL: Z99109; CAB12840.1; -.
DR FIR: B69818; B69818.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 118 AA; 13066 MW; 0F599F7011DD7F46 CRC64;

Query Match 12.8%; Score 81; DB 2; Length 118;
Best Local Similarity 26.2%; Pred. No. 56;
Matches 22; Conservative 22; Mismatches 24; Indels 16; Gaps 3;

QY 39 GQALKLNDWSDVSTSTFSKLRQLGPGVPTQSFWDNLEKETEGRL--QEMSKDLEEVKAK 96
DQ ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 30 GQALREKMTNVDSEETIKRLKS-----DGLALKDQLKAKESTDVIKDVGGE 79
QY 97 QPYLDDFOKKQOEEMELYRKQVE 120
DQ : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 80 LQTSI----KKQOEIKPHQODLQ 99

RESULT 12
QY Q7P280 PRELIMINARY; PRT; 96 AA.
AC Q7P280;
DT 01-MAR-2004 (TremBLrel. 26, Created)
DT 01-MAR-2004 (TremBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TremBLrel. 26, Last annotation update)
DE Hypothetical cytosolic protein.
GN Name=FNW007;
OS Fusobacterium nucleatum subsp. vincentii ATCC 49256.
OC Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;
OC Fusobacterium
OX NCBI_TaxID=209882;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 49256;
RA Karpal V., Ivanova N., Anderson I., Reznik G., Bhattacharyya A.,
RA Gardner W.L., Mikhailova N., Larsen N., D'Souza M., Walunas T.,
RA Haselkorn R., Overbeek R., Kyrides N.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL: AABF01000220; EAA23153.1; -.
DR InterPro: IPR004238; LEA.
DR Pfam: PF02987; LEA_4; 1.
KW Hypothetical protein.
SQ SEQUENCE 96 AA; 10436 MW; 7568787BAF86350 CRC64;

Query Match 12.3%; Score 77.5; DB 2; Length 96;
Best Local Similarity 24.2%; Pred. No. 85;
Matches 23; Conservative 20; Mismatches 47; Indels 5; Gaps 1;

QY 13 DLATVVDVLKSGRDYVSQFSGALGQNLKLLDNDWSDVSTSTFSKLRQLGPGVPTQSF 72
DQ : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 5 DDLTGILDEVKDTVDADKAKELKDEAVTKVGEUK-----DKATEKAGELKDKVVDKAKELK 59
QY 73 DNLEKETEGRLQEMSKDLEEVKAKVQPYLDDFOKK 107
DQ : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 60 EGAESKTSLEKDKAAEKAELKDKI SEGADSLINK 94

RESULT 13
QY Q9CHE9 PRELIMINARY; PRT; 79 AA.
AC Q9CHE9;
DT 01-JUN-2001 (TremBLrel. 17, Created)
DT 01-JUN-2001 (TremBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Hypothetical protein yhjA.
```

```
GN Name=yhjA; OrderedLocusNames=LL0783;
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IL1403;
RX MEDLINE=21235186; PubMed=11337471;
RA Bolotin A., Wincker P., Mauer S., Jaillon O., Malarne K.,
RA Weissenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
RL lactis ssp. lactis IL1403."
RL Genome Res. 11:731-753(2001).
DR EMBL: AE006312; AAK04881.1; -.
DR FIR: G86722; G86722.
DR InterPro: IPR008462; Cbbd.
DR Pfam: PF05532; Cbbd; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 79 AA; 8550 MW; 5EC62D428E5F0D53 CRC64;

Query Match 12.1%; Score 76.5; DB 2; Length 79;
Best Local Similarity 29.1%; Pred. No. 82;
Matches 23; Conservative 16; Mismatches 37; Indels 3; Gaps 1;

QY 42 LNLKLLDNDWSDVSTSTFSKLRQLGPGVPTQSFWDNLEKETEGRLQEMSKDLEEVKAKVQPYL 101
DQ ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 3 LNDKLDATKQVS---GKVKETTGVTGDEKLEAKGKTGLMGKAKEGLENIKDKASDLA 59
QY 102 DDFQKKQOEEMELYRKQVE 120
DQ : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 60 EDVAEKFNDDTVDSVRKHNE 78

RESULT 14
QY Q9PSS3 PRELIMINARY; PRT; 99 AA.
AC Q9PSS3;
DT 01-MAY-2000 (TremBLrel. 13, Created)
DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TremBLrel. 24, Last annotation update)
DE SNON protein (Fragments).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93293901; PubMed=8514802;
RA Nagase T., Nomura N., Ishii S.;
RT "Complex formation between proteins encoded by the ski gene family.";
RL J. Biol. Chem. 268:13710-13716(1993).
DR FIR: B46598; B46598.
FT NON_TER 1 1
FT NON_TER 56 57
FT NON_TER 99 99
SQ SEQUENCE 99 AA; 12042 MW; D39B07CA05BF176F CRC64;

Query Match 12.0%; Score 76; DB 2; Length 99;
Best Local Similarity 31.0%; Pred. No. 1.2e+02;
Matches 27; Conservative 20; Mismatches 20; Indels 20; Gaps 5;

QY 41 QLNKLLDNDWSDVSTSTFSKLRQLGPGVPTQSFWDNLEKETEGRLQEMSKDLEEVKAKVQPY 100
DQ : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 17 QMEVKMLSS-----SKSMKEL---TEE-QQNLOKLESLQNBHAQRMBESQLAELROR 64

QY 101 LD-----DFQKKQOEEMELYRKQVE 120
DQ || : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 65 LDHAEADRFQELQDELQRQEEA-RQKLE 90

RESULT 15
O68959
```

```

ID O68959 PRELIMINARY; PRT; 118 AA.
AC O68959;
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE M protein (Fragment).
GN Name=enm;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=13114;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ST448;
RA Beall B.W.;
RL Submitted (APR-1998) to the EMBL/GenBank/DDJ databases.
DR EMBL; AF058802; AAC14134.1; -.
FT NON_TER 1
FT NON_TER 118
SQ SEQUENCE 118 AA; 13597 MW; 744443E4DD757559 CRC64;

Query Match 12.0%; Score 76; DB 2; Length 118;
Best Local Similarity 29.0%; Pred. No. 1.4e+02;
Matches 31; Conservative 24; Mismatches 36; Indels 16; Gaps 5;

QY 17 VYVDVLKDSGRDYVSQFEGSALGKQLNLKLLDNWDSVTSTFSKLRQLGPTQBFWDNLE 76
Db 5 VAVAVL---GAGFANQTEYKAAEKKEVA-----DSNASSVAKLYNQIADLTDKNGEYLE 56

QY 77 KETGLRQEMSKDLEEVKAKVPYLDLDDFOKKWQEE---ELYRQKVE 120
Db 57 RIEE--LEERQKNLEKLERQSQVAAD---KHYYEQVKKHQYKQEQE 98

```

Search completed: December 21, 2004, 13:34:43
Job time : 192 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2004, 13:20:25 ; Search time 157 Seconds
(without alignments)
274.188 Million cell updates/sec

Title: US-09-803-918a-2_COPY_25_144

Perfect score: 631
Sequence: 1 DRFPQSPWDRVKDLATVYVD.....LDDFOKKWOEMELYRQKVE 120

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 1197155

Minimum DB seq length: 0
Maximum DB seq length: 120

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	476.5	75.5	120	4 AAU30469	Aau30469 Novel hum
2	476.5	75.5	120	4 AAU30267	Aau30267 Novel hum
3	410	65.0	119	4 AAU30468	Aau30468 Novel hum
4	197	31.2	65	6 ABP76131	Abp76131 Human GEN
5	193	30.6	42	2 AAR20164	Aar20164 Apo AI po
6	162	25.7	64	5 ABP31744	Abp31744 Human apo
7	138	21.9	32	2 AAR20165	Aar20165 Apo AI po
8	137	21.7	26	7 ADE76474	Ade76474 ApoAI G*
9	123	19.5	85	8 ADP87441	Adp87441 Monkey ap
10	119	18.9	85	8 ADP87442	Adp87442 Human apo
11	119	18.9	85	8 ADP87440	Adp87440 Baboon ap
12	118	18.7	22	1 AAP92072	Aap92072 Apolipop
13	116	18.4	21	1 AAP90956	Aap90956 Apo AI ep
14	108	17.1	84	2 AAY42554	Aay42554 Apolipop
15	107	17.0	85	8 ADP87439	Adp87439 Pig apoli
16	104	16.5	26	4 AAG52609	Aag52609 Apolipop
17	101.5	16.1	107	8 ADP87432	Adp87432 Rat apoli
18	101.5	16.1	107	8 ADP87431	Adp87431 Rat apoli
19	101	16.0	85	8 ADP87438	Adp87438 Chicken a
20	97	15.4	19	1 AAP92073	Aap92073 Apolipop
21	96.5	15.3	95	6 ABP70773	Abp70773 Human apo
22	94	14.9	30	7 ADC29665	Adc29665 Antioxi
23	94	14.9	30	7 ADC29666	Adc29666 Antioxi
24	92	14.6	18	7 ADC29658	Adc29658 Antioxi
25	91	14.4	16	2 AAR72706	Aar72706 Human apo

26	91	14.4	16	2 AAR34033	Aar34033 Sequence
27	91	14.4	16	2 AAY27065	Aay27065 Peptide S
28	89	14.1	16	6 ABP57175	Abp57175 Breast ca
29	89	14.1	30	7 ADC29668	Adc29668 Antioxi
30	89	14.1	30	7 ADC29664	Adc29664 Antioxi
31	87	13.8	17	8 ADH35704	Adh35704 Apolipop
32	87	13.8	17	8 ADH35715	Adh35715 Apolipop
33	87	13.8	18	7 ADC29657	Adc29657 Antioxi
34	87	13.8	18	7 ADC29660	Adc29660 Antioxi
35	86	13.6	30	7 ADC29667	Adc29667 Antioxi
36	86	13.6	30	7 ADC29671	Adc29671 Antioxi
37	85	13.5	30	7 ADC29669	Adc29669 Antioxi
38	85	13.5	30	7 ADC29670	Adc29670 Antioxi
39	84	13.3	18	7 ADC29663	Adc29663 Antioxi
40	84	13.3	18	7 ADC29659	Adc29659 Antioxi
41	84	13.3	96	8 ADP87419	Adp87419 Tupala gl
42	84	13.3	107	8 ADP87437	Adp87437 Mouse apo
43	84	13.3	107	8 ADP87433	Adp87433 Mouse apo
44	84	13.3	107	8 ADP87436	Adp87436 Mouse apo
45	84	13.3	107	8 ADP87434	Adp87434 Mouse apo

ALIGNMENTS

RESULT 1
AAU30469
ID AAU30469 standard; protein; 120 AA.
XX
AC AAU30469;
XX
DT 18-DEC-2001 (first entry)
XX
DE Novel human secreted protein #960.
XX
KW Human; vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX
OS Homo sapiens.
XX
PN WO200179449-A2.
XX
PD 25-OCT-2001.
XX
PF 16-APR-2001; 2001WO-US008656.
XX
PR 18-APR-2000; 2000US-00552929.
PR 26-JAN-2001; 2001US-00770160.
(HYSE-) HYSEQ INC.
PI Tang YT, Liu C, Drmanac RT;
WPI; 2001-611725/70.
XX
PT Nucleic acids encoding a range of human polypeptides, useful in genetic
PT vaccination, testing and therapy.
PS Claim 20; Page 297; 765pp; English.
CC The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising the
CC nucleic acids encoding the polypeptides and cells genetically engineered
CC to express them are also useful for producing the proteins. The proteins
CC are useful in genetic vaccination, testing and therapy, and can be used
CC as nutritional supplements. They may be used to increase stem cell

Query Match 65.0%; Score 410; DB 4; Length 119;
Best Local Similarity 100.0%; Pred. No. 3.9e-30;
Matches 78; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 NLKLLDNWDSVTSTFESKLRQLGVPTQSFWDNLEKETEGLRQEMSKDLSEVKAKVQPYLD 102
|||||
DB 2 NLKLLDNWDSVTSTFESKLRQLGVPTQSFWDNLEKETEGLRQEMSKDLSEVKAKVQPYLD 61
|||||

QY 103 DFQKKQWQEMELYRQKVE 120
|||||
DB 62 DFQKKQWQEMELYRQKVE 79
|||||

RESULT 4

ABP76131
ID ABP76131 standard; protein; 65 AA.

AC ABP76131;

XX 21-FEB-2003 (first entry)

XX Human GENSET protein SEQ ID 457.

XX Cytostatic; antiinflammatory; nootropic; neuroprotective; cardiant;
KW gastrointestinal; gene therapy; GENSET; heavy metal toxicity; cancer;
KW inflammatory disease; immune disorder; neuromuscular; toxicity;
KW central nervous system; cardiovascular; gastrointestinal.

XX Homo sapiens.

XX WO200283898-A1.

XX 24-OCT-2002.

XX 18-APR-2001; 2001WO-IB000914.

XX 18-APR-2001; 2001WO-IB000914.

XX (GEST) GENSET.

XX Bejanin S, Tanaka H, Dumas Milne Edwards J, Jobert S, Giordano J;
PI WPI; 2003-075548/07.

XX New GENSET polynucleotides and polypeptides, useful for treating heavy
PT metal toxicity, cancer, inflammatory diseases, immune disorders, and the
PT neuromuscular, CNS, cardiovascular or gastrointestinal effects of the
PT toxicity.

XX Claim 14; Page 535; 735pp; English.

XX The present invention relates to novel GENSET polynucleotides (ABZ36404-
CC ABZ36911) encoding polypeptides (ABP75963-ABP76368). The polynucleotides
CC and polypeptides are useful in screening and diagnostic assays for
CC abnormal GENSET expression and/or biological activity. They are also
CC useful for screening of compounds for treating or preventing GENSET-
CC related disorders, such as heavy metal toxicity, cancer, inflammatory
CC diseases, immune disorders, and the neuromuscular, central nervous system
CC (CNS), cardiovascular or gastrointestinal effects of the toxicity

XX Sequence 65 AA;

Query Match 31.2%; Score 197; DB 6; Length 65;
Best Local Similarity 100.0%; Pred. No. 9.6e-11;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DEPPQSPWDRVKDLATVYVDVLKDSGRDYVSQFEGSA 37
|||||

DB 25 DEPPQSPWDRVKDLATVYVDVLKDSGRDYVSQFEGSA 61
|||||

RESULT 5

AAR20164

ID AAR20164 standard; peptide; 42 AA.

XX AAR20164;

XX 25-MAR-2003 (revised)

DT 01-APR-1992 (first entry)

XX Apo AI polypeptide (A).

XX Monoclonal antibody; MAB AI-11; epitope; diagnosis.

XX Synthetic.

XX Key Location/Qualifiers

FT Peptide 1..28

FT Peptide 2..28

FT Peptide 7..28

FT Peptide 10..28

FT Peptide 11..42

FT Peptide 11..28

FT Peptide 13..28

XX WO9118619-A.

XX 12-DEC-1991.

XX 07-JUN-1990; 90US-00534761.

XX 07-JUN-1990; 90US-00534761.

XX (SCRI) SCRIPPS CLINIC & RES.

XX Curtiss LK, Banka CL, Bonnet DJ, Smith RS;

XX WPI; 1992-007201/01.

XX New Apo AI polypeptide(s) and anti-Apo antibodies - used to detect Apo AI
PT in vascular fluid samples and increase LCAT-mediated cholesterol
PT esterification in humans.

XX Claim 2; Page 68; 87pp; English.

XX The peptides represented in AAR20164-65 are capable of immunologically
CC mimicking an Apo AI epitope. They are useful in diagnosis and detection
CC of Apo AI in fluids and for preparing anti-Apo AI antibodies. It is
CC useful in therapeutic methods for increasing LCAT-mediated cholesterol
CC esterification in humans. The Apo AI polypeptide is selected from the
CC peptides indicated in the features and includes amino acids 13-28
CC defining a conserved native epitope on Apo AI capable of immunoreacting
CC with monoclonal antibody MAB AI-11. (Updated on 25-MAR-2003 to correct PA
CC field.)

XX Sequence 42 AA;

Query Match 30.6%; Score 193; DB 2; Length 42;
Best Local Similarity 100.0%; Pred. No. 1.3e-10;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 QEMSKDLSEVKAKVQPYLDDFQKKQWQEMELYRQKVE 120
|||||

DB 1 QEMSKDLSEVKAKVQPYLDDFQKKQWQEMELYRQKVE 37
|||||

RESULT 6

ABP31744

ID ABP31744 standard; protein; 64 AA.

XX ABP31744;

XX 08-JUL-2002 (first entry)

XX Human apolipoprotein-like ORF717 protein, SEQ ID NO:1434.

XX

KW Human; ORF; open reading frame; ORFX; drug screening; diagnosis;
KW disease monitoring; cytokine; cell proliferation; cell differentiation;
KW immune modulation; haematopoiesis regulation; tissue growth;
KW angiogenesis; activin; inhibin; chemotactic; chemokinetic; haemostatic;
KW thrombolytic; tumour inhibition; bodily characteristics; fertility;
KW behaviour; cancer; proliferative disorder; neurological disorder;
KW cardiovascular disease; immune system disorder; organ transplantation;
KW tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
KW hypothyroidism; cholesterol ester storage disease; infection; vulnery;
KW vasotropic; antipsoriatic; antidiabetic; cytostatic; neurotropic;
KW neuroprotective; antiatherosclerotic; anticoagulant; thrombolytic;
KW cardiant; hypotensive; antithyroid; antiinflammatory; immunomodulator;
KW dermatological; analgesic; virucide; antibacterial; fungicide.
XX
OS Homo sapiens.
XX
PN WO200190366-A2.
XX
XX
PD 29-NOV-2001.
XX
XX
PF 24-MAY-2001; 2001WO-US017076.
XX
XX
PR 24-MAY-2000; 2000US-0206690P.
XX
XX
PA (CURA-) CURAGEN CORP.
XX
PI Leach MD, Shinkets RA;
XX
XX WPI; 2002-106200/14.
DR N-PSDB; ABN75770.
XX
XX
PT Novel human polypeptides and polynucleotides useful for diagnosing,
PT preventing and treating cardiovascular disease, neurodegenerative,
PT hyperproliferative disorders and disorders related to organ
PT transplantation.
XX
XX Claim 10; Page 625; 2508pp; English.
XX
XX Sequences ABP31028-ABP35561 represent 4534 novel human proteins
CC designated ORF (open reading frame) 1-4534, and sequences ABN75054-
CC ABN79587 represent cDNAs encoding them. The invention also encompasses
CC polypeptides at least 80% identical to the ORF1-ORF4534 (collectively
CC referred to as ORFX) proteins, polynucleotides at least 85% identical to
CC the ORFX nucleic acid sequences, vectors and host cells comprising ORFX
CC polynucleotides, the recombinant production of ORFX proteins, antibodies
CC specific for ORFX proteins, methods of detecting ORFX polynucleotides and
CC polypeptides, methods of screening for modulators of ORFX expression or
CC activity, and methods of screening individuals for a predisposition to an
CC ORFX-associated disorder. The ORFX proteins of the invention have a wide
CC range of biological activities, such as cytokine, cell proliferation,
CC cell differentiation, immune modulation, haematopoiesis regulation,
CC tissue growth, angiogenesis, activin or inhibin activity, chemotactic/
CC chemokinetic activity, haemostatic activity, thrombolytic activity,
CC receptor/ligand, antiinflammatory activity, tumour inhibition activity,
CC and antiinfective activity, and may also be involved in the determination
CC of bodily characteristics, fertility and behaviour. ORFX proteins,
CC nucleic acids and antibodies may be used in the treatment of cancers,
CC other proliferative disorders such as psoriasis and benign tumours,
CC neurological disorders such as epilepsy and Alzheimer's disease,
CC cardiovascular diseases, immune system disorders, disorders related to
CC organ transplantation, disorders of tissue growth and regeneration,
CC diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester
CC storage disease, and infectious diseases caused by viral, bacterial,
CC fungal and other pathogens. ORFX nucleic acids may also be used as a
CC source of primers and probes, in the detection of ORFX genomic sequences
CC or transcripts, in the identification and cloning of homologous
CC sequences, in genetic diagnosis, and in forensic biology. The ORFX
CC nucleic acids may additionally be used to produce transgenic animals
CC which may be useful for studying the function and/or activity of ORFX
CC protein, and in drug screening. The ORFX proteins may also be used as
CC immunogens to generate specific antibodies, which are useful in the
CC diagnosis, treatment and monitoring of ORFX-associated diseases
XX

SQ Sequence 64 AA;
Query Match 25.7%; Score 162; DB 5; Length 64;
Best Local Similarity 75.0%; Pred. No. 1.6e-07;
Matches 30; Conservative 4; Mismatches 6; Indels 0; Gaps 0;
QY 2 EPPOSPWRVVDLATVYVDVLKDSGRDVSQFEGSALGKQ 41
DB 25 DEPOSQMDRVKDFATVYVYDAVKDSGRNVYSQFESLTGQQ 64
RESULT 7
AAR20165
ID AAR20165 standard; protein; 32 AA.
XX
AC AAR20165;
XX
XX 25-MAR-2003 (revised)
DT 01-APR-1992 (first entry)
XX
XX Apo AI polypeptide (B).
XX Monoclonal antibody; MAB AI-14; epitope; diagnosis.
XX
XX Synthetic.
XX
FH Key Location/Qualifiers
FT Peptide 1..32
FT Peptide 1..21
FT Peptide 5..28
FT Peptide 5..21
FT Peptide 6..28
FT Peptide 6..21
FT Misc-difference 11 /label= GLU, PHE
XX
XX WO9118619-A.
XX
XX 12-DEC-1991.
PD
PF 07-JUN-1990; 90US-00534761.
PR 07-JUN-1990; 90US-00534761.
XX (SCRI) SCRIPPS CLINIC & RES.
XX
XX Curtiss LK, Banka CL, Bonnet DJ, Smith RS;
XX WPI; 1992-007201/01.
XX
XX New Apo AI polypeptide(s) and anti-Apo antibodies - used to detect Apo AI
FT in vascular fluid samples and increase LCAT-mediated cholesterol
PT esterification in humans.
XX
XX Claim 5; Page 68; 87pp; English.
XX
XX The peptides represented in AAR20164-65 are capable of immunologically
CC mimicking an Apo AI epitope. They are useful in diagnosis and detection
CC of Apo AI in fluids and for preparing anti-Apo AI antibodies. It is
CC useful in therapeutic methods for increasing LCAT-mediated cholesterol
CC esterification in humans. The Apo AI polypeptide is selected from the
CC peptides indicated in the features and includes amino acids 6-21 defining
CC a conserved native epitope on Apo AI capable of immunoreacting with
CC monoclonal antibody MAB AI-4. (Updated on 25-MAR-2003 to correct PA
CC field.)
XX
XX Sequence 32 AA;
Query Match 21.9%; Score 138; DB 2; Length 32;
Best Local Similarity 96.3%; Pred. No. 1.2e-05;
Matches 26; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 94 KAKVQPYLDDFOKQWQEMELYRQKVE 120

Db	1	KAKVQPYLDDXQKKWQEMELYRQKVE	27
RESULT 8			
ADP76474			
ID	ADP76474	standard; peptide; 26 AA.	
XX	AC	ADP76474;	
XX	DT	29-JAN-2004 (first entry)	
XX	DE	ApoAI G* amphipathic helical domain-derived peptide - SEQ ID 21.	
XX	KW	G-type; atherosclerosis; inflammatory; G* amphipathic helix;	
XX	KW	antiarteriosclerotic; antirheumatic; antiarthritic; immunosuppressive;	
XX	KW	dermatological; antiinflammatory; osteopathic; neuroprotective;	
XX	KW	neurotropic; anti-HIV; virucide; antibacterial; rheumatoid arthritis;	
XX	KW	systemic lupus erythematosus; polyarteritis nodosa; osteoporosis;	
XX	KW	Alzheimer's disease; AIDS; infection; apoAI.	
XX	OS	Unidentified.	
XX	PH	Key	Location/Qualifiers
FT	Misc-difference	1. .26	
FT	/note=	"Optionally D-form residue"	
FT	Modified-site	1	
FT	/note=	"Optional N-terminal acetyl"	
FT	Modified-site	26	
FT	/note=	"Optional C-terminal amide"	
XX	US2003191057-A1.		
PN	09-OCT-2003.		
XX	05-APR-2002; 2002US-00120508.		
XX	05-APR-2002; 2002US-00120508.		
XX	(REGC) UNIV CALIFORNIA.		
PA	Fogelman AM, Navab M;		
PI	WPI; 2003-831662/77.		
XX	New G-type polypeptides for preventing or ameliorating symptoms of		
XX	atherosclerosis or pathologies associated with an inflammatory responses		
XX	e.g. rheumatoid arthritis, osteoporosis, Alzheimer's disease, AIDS or		
XX	infections.		
XX	Claim 7; SEQ ID NO 21; 29pp; English.		
PS	The invention relates to a novel isolated G-type polypeptide that		
XX	ameliorates a symptom of atherosclerosis or other pathology associated		
XX	with an inflammatory response comprising an amphipathic helical peptide		
XX	having charged residues on the polar face of the peptide and possessing a		
XX	wide non-polar face. The polypeptide of the invention demonstrates		
XX	antiarteriosclerotic, antirheumatic, antiarthritic, immunosuppressive,		
XX	dermatological, antiinflammatory, osteopathic, neuroprotective,		
XX	neurotropic, anti-HIV, virucide and antibacterial activities. The		
XX	composition and methods of the invention may be useful in preventing or		
XX	ameliorating one or more symptoms of atherosclerosis and/or pathologies		
XX	characterised by an inflammatory response such as rheumatoid arthritis,		
XX	systemic lupus erythematosus, polyarteritis nodosa, osteoporosis,		
XX	Alzheimer's disease, AIDS or infection. The current sequence is that of		
XX	the apoAI G* amphipathic helical domain-derived peptide of the invention.		
XX	Sequence 26 AA;		
XX	Query Match	21.7%; Score 137; DB 7; Length 26;	
XX	Best Local Similarity	100.0%; Pred. No. 1.1e-05;	
XX	Matches	26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	8	WDRVKDLATVYVDVLKDSGRDYVSQF	33
Db	1	WDRVKDLATVYVDVLKDSGRDYVSQF	26
RESULT 9			
ADP87441			
ID	ADP87441	standard; protein; 85 AA.	
XX	AC	ADP87441;	
XX	DT	09-SEP-2004 (first entry)	
XX	DE	Monkey apolipoprotein A (APO-A) precursor protein.	
XX	KW	Browseable database system; ontology; protein analysis;	
XX	KW	gene product classification; genomic analysis; apolipoprotein A; APO-A;	
XX	KW	monkey.	
XX	OS	Macaca fascicularis.	
XX	PN	WO2004053769-A2.	
XX	24-JUN-2004.		
XX	09-DEC-2003; 2003WO-US038935.		
XX	09-DEC-2002; 2002US-0431879P.		
XX	(APPL-) APPLERA CORP.		
XX	Thomas PD, Kejariwal A, Campbell MJ, Mi H, Diemer K, Guo N;		
PI	Ladunga I, Lazareva B, Muruganujan A, Rabkin S, Vandergriff J;		
PI	Doremieux O;		
XX	WPI; 2004-480967/45.		
DR	Browseable database system, for e.g. analyzing protein sequences, and		
PT	predicting a biological role, comprises a database, an ontology of		
PT	categories of biological functions, an input receptive, a recognizer, and		
PT	an output.		
XX	Disclosure; SEQ ID NO 48; 113pp; English.		
PS	The invention relates to a browsable database system for use with		
XX	biological information, comprises a database, an ontology of categories		
XX	of biological functions, an input receptive, a recognizer and an output.		
XX	The invention also provides a method of operation for use with a		
XX	browseable biological database system. The browsable database system and		
XX	methods are useful in analysing protein sequences, classifying gene		
XX	products, predicting biological role for pathway building, enhancing		
XX	interpretation of expression information, providing protein function and		
XX	in facilitating comparative genomic analysis. The present sequence is		
XX	monkey apolipoprotein A (APO-A) IV precursor protein. This sequence is		
XX	used to illustrate the method of the invention.		
XX	Sequence 85 AA;		
XX	Query Match	19.5%; Score 123; DB 8; Length 85;	
XX	Best Local Similarity	31.8%; Pred. No. 0.00095;	
XX	Matches	27; Conservative 23; Mismatches 35; Indels 0; Gaps 0;	
QY	36	SALGKQLNLKLDNDNSTFTSKLREQLGPTVQRFWDNLKXETEGLRQEMSKDLREKVA	95
Db	1	SELTQQLNALFDKLGENVNTYAGDIQKLVPPATELHERLAKNDSEKKEIRKELEBVA	60
QY	96	KVQPYLDDDFQKKWQEMELYRQKVE	120
Db	61	RLLPANEVSQKIGENVRELQRL	85
RESULT 10			
ADP87442			

ID	ADP87442	standard; protein; 85 AA.
XX	AC	
XX	ADP87442;	
DT	09-SEP-2004	(first entry)
XX	Human apolipoprotein A (APO-A)	precursor protein.
XX	Browsable database system; ontology; protein analysis;	
KW	gene product classification; genomic analysis; apolipoprotein A; APO-A;	
KW	human.	
XX	Homo sapiens.	
OS	WO2004053769-A2.	
XX	24-JUN-2004.	
XX	09-DEC-2003; 2003WO-US038935.	
XX	09-DEC-2002; 2002US-0431879P.	
XX	(APPL-) APPLERA CORP.	
PA	Thomas PD, Kejariwal A, Campbell MJ, Mi H, Diemer K, Guo N;	
XX	Ladunga I, Lazareva B, Muruganujan A, Rabkin S, Vandergriff J;	
PI	Doremieux O;	
PI	WPI; 2004-480967/45.	
DR	Browsable database system, for e.g. analyzing protein sequences, and	
XX	predicting a biological role, comprises a database, an ontology of	
PT	categories of biological functions, an input receptive, a recognizer, and	
PT	an output.	
XX	Disclosure; SEQ ID NO 49; 113pp; English.	
XX	The invention relates to a browsable database system for use with	
CC	biological information, comprises a database, an ontology of categories	
CC	of biological functions, an input receptive, a recognizer and an output.	
CC	The invention also provides a method of operation for use with a	
CC	browsable biological database system. The browsable database system and	
CC	methods are useful in analysing protein sequences, classifying gene	
CC	products, predicting biological role for pathway building, enhancing	
CC	interpretation of expression information, providing protein function and	
CC	in facilitating comparative genomic analysis. The present sequence is	
CC	human apolipoprotein A (APO-A) IV precursor protein. This sequence is	
CC	used to illustrate the method of the invention.	
XX	Sequence 85 AA;	
XX	Query Match 18.9%; Score 119; DB 8; Length 85;	
XX	Best Local Similarity 29.4%; Pred. No. 0.0022;	
XX	Matches 25; Conservative 25; Mismatches 35; Indels 0; Gaps 0	
QY	36	SALGKGLNLKLDNDWSDVTSTFSKLRLEQLGPTQTFWDNLKTEGRLQRMKDLREKVA 95
Db	1	SELTQGNALFQDKLGEVNTVAGDQLQKLVFPFATELHRLAKUSEKLEIGKELELRA 60
QY	96	KVQPYLDDFKKQWEMELYRKQVE 120
Db	61	RLLPHANEVSQKIGDNLRELQORLE 85
RESULT 11		
ADP87440		
ID	ADP87440	standard; protein; 85 AA.
XX	ADP87440;	
XX	09-SEP-2004	(first entry)
DT	Baboon apolipoprotein A (APO-A)	IV precursor protein.
XX		

XX	Browsable database system; ontology; protein analysis;
KW	gene product classification; genomic analysis; apolipoprotein A; APO-A;
KW	baboon.
XX	
OS	Papio anubis.
XX	
FN	WO2004053769-A2.
XX	
PD	24-JUN-2004.
XX	
FF	09-DEC-2003; 2003WO-US038935.
XX	
PR	09-DEC-2002; 2002US-0431879P.
XX	
PA	(APPL-) APPLERA CORP.
XX	
PI	Thomas PD, Kejariwal A, Campbell MJ, Mi H, Diemer K, Guo N;
PI	Ladunga I, Lazareva B, Muruganujan A, Rabkin S, Vandergriff J;
PI	Doremieux O;
XX	
DR	WPI; 2004-480967/45.
XX	
PT	Browsable database system, for e.g. analyzing protein sequences, and
PT	predicting a biological role, comprises a database, an ontology of
PT	categories of biological functions, an input receptive, a recognizer, and
PT	an output.
XX	
PS	Disclosure; SEQ ID NO 47; 113pp; English.
XX	
CC	The invention relates to a browsable database system for use with
CC	biological information, comprises a database, an ontology of categories
CC	of biological functions, an input receptive, a recognizer and an output.
CC	The invention also provides a method of operation for use with a
CC	browsable biological database system. The browsable database system and
CC	methods are useful in analysing protein sequences, classifying gene
CC	products, predicting biological role for pathway building, enhancing
CC	interpretation of expersion information, providing protein function and
CC	in facilitating comparative genomic analysis. The present sequence is
CC	baboon apolipoprotein A (APO-A) IV precursor protein. This sequence is
CC	used to illustrate the method of the invention.
XX	
SQ	Sequence 85 AA;
	Query Match 18.9%; Score 119; DB 8; Length 85;
	Best Local Similarity 30.6%; Pred No. 0.0022;
	Matches 26; Conservative 24; Mismatches 35; Indels 0; Gaps 0;
QY	36 SALGKQLNKLKLDNWDSTFTFSKLRQLGQVPTQFMDNLEKTEGLRQEMSKOLEEYKA 95
Db	1 SELTQQLNALFDQKLGEVNTVAGDLOKKLVFFATELHERLAKDSKKLEEIRKEELEEVRA 60
QY	96 KVQPYLDDFQKKQOEEMELRYQKVE 120
Db	61 RLLPHANEVSKIGENVRELQQRLE 85
RESULT 12	
AAP92072	
ID	AAP92072 standard; peptide; 22 AA.
XX	
AC	AAP92072;
XX	
DT	25-MAR-2003 (revised)
DT	31-OCT-2002 (revised)
DT	10-APR-1990 (first entry)
XX	
DE	Apolipoprotein AI (Apo AI) polypeptide.
XX	
KW	Apolipoprotein AI; Apo AI; high-density lipoprotein; HDL;
KW	coronary artery disease; CAD; Apo AI epitope.
XX	
OS	Homo sapiens.

XX FH Key Location/Qualifiers
 FT Peptide 1. .22 /label= AI90-111
 FT Peptide 1. .22 /label= AI90-111
 FT Peptide 1. .16 /label= AI90-105
 FT Peptide 4. .12 /label= AI93-101
 FT Peptide 6. .16 /label= AI95-105
 FT Peptide 7. .12 /label= AI96-101
 FT Peptide 11. .16 /label= AI100-105
 FT Peptide 12. .22 /label= AI101-111

XX WO8904486-A.

XX 18-MAY-1989.

XX 02-NOV-1988; 88WO-US003903.

XX 03-NOV-1987; 87US-00116248.

XX (SCRI) SCRIPPS CLINIC & RES FOUND.

XX Curtiss LK, Smith RR;

XX WPI; 1989-165740/22.

XX New monoclonal antibody and polypeptide antigens - directed against APO AI-HDL epitope, useful in diagnostic assays.

XX Claim 2; Table 1-2; 62pp; English.

XX AI90-111 and smaller peptides contained within this sequence. AI95-105, AI190-105 and AI190-111 are specifically claimed. Such Apo AI peptides are capable of immunologically mimicking a native conserved Apo AI epitope. They may be helpful in the diagnosis of risk of coronary heart disease. (Updated on 31-OCT-2002 to add missing OS field.) (Updated on 25-MAR-2003 to correct PR field.) (Updated on 25-MAR-2003 to correct PA field.)

XX Sequence 22 AA;

Query Match 18.7%; Score 118; DB 1; Length 22;
 Best Local Similarity 100.0%; Pred. No. 0.00052;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 90 LEEVKAKVQPYLDDFOKKWQEE 111
 DB 1 LEEVKAKVQPYLDDFOKKWQEE 22

RESULT 13

AAP90956

ID AAP90956 standard; peptide; 21 AA.

XX AAP90956;

XX 25-MAR-2003 (revised)

DT 19-FEB-1990 (first entry)

XX Apo AI epitope.

XX High density lipoproteins.

XX Synthetic.

XX WO8909403-A.

XX 05-OCT-1989.
 XX 27-MAR-1989; 89WO-US001262.
 XX 29-MAR-1988; 88US-00174698.
 XX (SCRI) SCRIPPS CLINIC & RES FOUND.
 XX Curtiss LK, Smith RS;
 XX WPI; 1989-309644/42.
 XX Apo lipoprotein peptide and monoclonal antibody against it - used to assay peptide in vascular fluid to identify risk of coronary artery disease.
 XX Claim 2; Page 41; 49pp; English.
 XX Apo AI contains the sequence below (Claim 1 claims AAs 1-15). There is an inverse correlation between plasma levels of Apo AI and coronary artery disease risk. The monoclonal antibody recognises this epitope on native Apo AI. (Updated on 25-MAR-2003 to correct PA field.)

XX Sequence 21 AA;

Query Match 18.4%; Score 116; DB 1; Length 21;
 Best Local Similarity 100.0%; Pred. No. 0.00075;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DEPPQSPWDRVKDLATVYVDV 21

DB 1 DEPPQSPWDRVKDLATVYVDV 21

RESULT 14

AAY42554

ID AAY42554 standard; peptide; 84 AA.

XX AAY42554;

XX 20-DEC-1999 (first entry)

XX Apolipoprotein A-IV derived lipid oxidation suppressant peptide #2.
 XX Cholesterol; cardiovascular disease; heart disease; atherosclerosis;
 XX lipoprotein; angina; myocardial infarction; stroke; thrombosis;
 XX antioxidant; hypolipidaemic; apolipoprotein.

XX Synthetic.
 XX Homo sapiens.

XX Key Location/Qualifiers
 FT Modified-site 84 /note= "Optionally has a C-terminal amide"

XX WO950286-A2.
 XX 07-OCT-1999.

XX 25-MAR-1999; 99WO-US006580.

XX 31-MAR-1998; 98US-0080131P.

XX (UYCI-) UNIV CINCINNATI.

XX Hui DY, Tso P;

XX WPI; 1999-580739/49.

XX Treating conditions associated with lipid oxidation or preventing oxidation in lipid-containing food, lipid containing pharmaceuticals or cosmetic or dermatological compositions.

```

XX Claim 4; Page 53; 73pp; English.
PS This sequence represents a peptide (#2) derived from apolipoprotein (apo)
XX A-IV with lipid oxidation inhibitory activity. Lipid oxidation plays a
CC role in the development of atherosclerosis, a main cause of coronary
CC heart disease. Atherosclerosis is thought to begin with local injury to
CC the arterial endothelium, followed by proliferation of arterial smooth
CC muscle cells, along with deposition of lipid and accumulation of foam
CC cells in the lesion. As the atherosclerotic plaque develops, it
CC progressively occludes more and more blood vessel and can eventually lead
CC to ischaemia or infarction. Both the plasma concentration and qualitative
CC characteristics of low density lipoproteins (LDL) are risk factors in
CC atherogenesis. Oxidation causes important changes in the primary
CC structure of the main LDL apolipoprotein, apolipoprotein B100 (apo B-
CC 100). These changes, by helping LDL absorption by macrophages, causes the
CC intracellular accumulation of esters of cholesterol and the formation of
CC foam cells, with subsequent development of the atherosclerotic plaque.
CC The apo A-IV derived peptides can be used for inhibiting lipid oxidation.
CC The peptides can be used for treating conditions associated with lipid
CC oxidation. For example, they can be used for treating or inhibiting the
CC progression of atherosclerosis. They can also be used for preventing
CC oxidation in lipid-containing foods, lipid-containing pharmaceuticals or
CC cosmetic or dermatological compositions. As these peptides comprise
CC specific portions of the native apo A-IV protein, there should be no
CC immunogenicity problems with their administration to humans
XX
SQ Sequence 84 AA;

Query Match 17.1%; Score 108; DB 2; Length 84;
Best Local Similarity 27.8%; Pred. No. 0.023;
Matches 25; Conservative 22; Mismatches 35; Indels 8; Gaps 1;

QY 8 WNRVKDLATVYDVVLKDSGRDYVSQFEGSALGKQLNKLDDNWDSVTSTFSLRQLGVPV 67
DB || : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2 WD-----YFTQLSNNAKEAVEQLQKTDVTQQLNTLFDQKLGNINTYADDLQNKLVFP 53

QY 68 TORFWDNLEKETGLRQEMSKDLEEVKAKV 97
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
54 AVOLSGHLTKETVRVREEIQKELEDURAMV 83

RESULT 15
ADP87439
ID ADP87439 standard; protein; 85 AA.
AC ADP87439;
XX
DT 09-SEP-2004 (first entry)
XX
DE Pig apolipoprotein A (APO-A) IV precursor protein.
XX
KW Browseable database system; ontology; protein analysis;
KW Gene product classification; genomic analysis; apolipoprotein A; APO-A;
KW pig.
XX
OS Sus scrofa.
XX
PN WO2004053769-A2.
XX
PD 24-JUN-2004.
XX
PF 09-DEC-2003; 2003WO-US038935.
XX
PR 09-DEC-2002; 2002US-0431879P.
XX
PA (APPL-) APPLERA CORP.
XX
PI Thomas PD, Kejarawal A, Campbell MJ, Mi H, Diemer K, Guo N;
PI Ladunga I, Lazareva B, Muruganujan A, Rabkin S, Vandergriff J;
PI Doremieux O;
XX
DR WPI; 2004-480967/45.

XX Browseable database system, for e.g. analyzing protein sequences, and
PT predicting a biological role, comprises a database, an ontology of
PT categories of biological functions, an input receptive, a recognizer, and
PT an output.
XX
PS Disclosure; SEQ ID NO 46; 113pp; English.
XX
CC The invention relates to a browsable database system for use with
CC biological information, comprises a database, a recognizer and an output.
CC The invention also provides a method of operation for use with a
CC browsable biological database system. The browsable database system and
CC methods are useful in analysing protein sequences, classifying gene
CC products, predicting biological role for pathway building, enhancing
CC interpretation of expression information, providing protein function and
CC in facilitating comparative genomic analysis. The present sequence is pig
CC apolipoprotein A (APO-A) IV precursor protein. This sequence is used to
XX illustrate the method of the invention.
XX
SQ Sequence 85 AA;

Query Match 17.0%; Score 107; DB 8; Length 85;
Best Local Similarity 27.4%; Pred. No. 0.029;
Matches 23; Conservative 25; Mismatches 36; Indels 0; Gaps 0;

QY 36 SALGKQLNKLDDNWDSVTSTFSLRQLGVPVTOFWDNLEKETGLRQEMSKDLEEVKKA 95
DB || : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1 SELTQQLNTLFDQKLGEVNTYTEDLQKLVPPATLHERLTKDSEKJKEETRRELELRA 60

QY 96 KVQPYLDDDFQKKWQEMELYRQKV 119
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
61 RLLPHATEVSVQKIGDNVRELQORL 84

Search completed: December 21, 2004, 13:37:26
Job time : 159 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2004, 12:34:05 ; Search time 142 Seconds
(without alignments)
302.396 Million cell updates/sec

Title: US-09-803-918a-2_COPY_25_144

Perfect score: 631
Sequence: 1 DEFPQSPWVRKDLATVYVD.....LDDFKKQGEMLYRQKVE 120

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1589859 segs, 357834939 residues

Total number of hits satisfying chosen parameters: 783582

Minimum DB seq length: 0
Maximum DB seq length: 120

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	387.5	61.4	93	15 US-10-038-854-403	Sequence 403, App
2	387.5	61.4	93	15 US-10-038-854-405	Sequence 405, App
3	318	50.4	86	15 US-10-424-599-276814	Sequence 276814,
4	227	36.0	43	15 US-10-465-789A-99	Sequence 89, Appl
5	162	25.7	64	11 US-09-864-408A-1434	Sequence 1434, Ap
6	137	21.7	26	14 US-10-120-508-21	Sequence 21, Appl
7	122	19.3	22	15 US-10-465-789A-50	Sequence 52, Appl
8	118	18.7	22	15 US-10-465-789A-50	Sequence 52, Appl
9	112	17.7	22	15 US-10-465-789A-49	Sequence 49, Appl
10	96.5	15.3	95	17 US-10-487-096-2	Sequence 2, Appl
11	94	14.9	30	14 US-10-142-238A-40	Sequence 40, Appl
12	94	14.9	30	14 US-10-142-238A-41	Sequence 41, Appl
13	92	14.6	18	14 US-10-142-238A-33	Sequence 33, Appl

14	89	14.1	16	17	US-10-700-340-37	Sequence 37, Appl
15	89	14.1	30	14	US-10-142-238A-39	Sequence 39, Appl
16	89	14.1	30	14	US-10-142-238A-43	Sequence 43, Appl
17	87	13.8	17	15	US-10-601-100-44	Sequence 44, Appl
18	87	13.8	18	14	US-10-142-238A-32	Sequence 32, Appl
19	87	13.8	18	14	US-10-142-238A-35	Sequence 35, Appl
20	86	13.6	30	14	US-10-142-238A-42	Sequence 42, Appl
21	86	13.6	30	14	US-10-142-238A-46	Sequence 46, Appl
22	85	13.5	30	14	US-10-142-238A-44	Sequence 44, Appl
23	85	13.5	30	14	US-10-142-238A-45	Sequence 45, Appl
24	84	13.3	18	14	US-10-142-238A-34	Sequence 34, Appl
25	84	13.3	18	14	US-10-142-238A-38	Sequence 38, Appl
26	83	13.2	18	14	US-10-142-238A-36	Sequence 36, Appl
27	83	13.2	18	14	US-10-142-238A-37	Sequence 37, Appl
28	77	12.2	90	15	US-10-038-854-224	Sequence 224, App
29	77	12.2	90	15	US-10-038-854-231	Sequence 231, App
30	76	12.0	107	14	US-10-023-066A-92	Sequence 92, Appl
31	74	11.7	14	15	US-10-601-100-43	Sequence 43, Appl
32	72	11.4	113	17	US-10-480-848A-5	Sequence 5, Appl
33	72	11.4	62	9	US-09-864-761-40424	Sequence 40424, A
34	68.5	10.9	92	16	US-10-767-701-61738	Sequence 61738, A
35	67.5	10.7	72	16	US-10-767-701-37365	Sequence 37365, A
36	67	10.6	21	14	US-10-076-047A-34	Sequence 34, Appl
37	66.5	10.5	105	17	US-10-639-067-135	Sequence 135, App
38	66	10.5	13	14	US-10-014-340-123	Sequence 123, App
39	66	10.5	13	15	US-10-264-309-76	Sequence 76, Appl
40	66	10.5	13	15	US-10-601-100-36	Sequence 36, Appl
41	66	10.5	105	9	US-09-767-215-10	Sequence 10, Appl
42	66	10.5	119	16	US-10-767-701-58043	Sequence 58043, A
43	65.5	10.4	99	15	US-10-674-755-16	Sequence 16, Appl
44	65.5	10.4	118	14	US-10-136-728-32	Sequence 32, Appl
45	65	10.3	76	14	US-10-106-698-6884	Sequence 6884, Ap

ALIGNMENTS

RESULT 1
US-10-038-854-403
; Sequence 403, Application US/10038854
; Publication No. US20040022781A1
; GENERAL INFORMATION:
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Li, Li
; APPLICANT: Wolenc, Adam R
; APPLICANT: Vernet, Corine
; APPLICANT: Eissen, Andrew J
; APPLICANT: Liu, Xiaohong
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Shimkets, Richard A
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spaderna, Steven K
; APPLICANT: Gorman, Linda
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Gangolli, Esha A
; APPLICANT: Guo, Xiaojia S
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Rastelli, Luca
; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc
; APPLICANT: Burgess, Catherine E
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gunther, Erik
; APPLICANT: Smithson, Glennnda
; APPLICANT: Millet, Isabelle
; APPLICANT: MacDougall, John R
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-230
; CURRENT APPLICATION NUMBER: US/10/038,854
; CURRENT FILING DATE: 2003-01-22

```
; PRIOR APPLICATION NUMBER: 60/258,928
; PRIOR FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: 60/259,415
; PRIOR FILING DATE: 2001-01-02
; PRIOR APPLICATION NUMBER: 60/259,785
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/269,814
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/279,832
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,833
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,863
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/283,889
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,447
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/286,583
; PRIOR FILING DATE: 2001-04-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 403
; LENGTH: 93
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-038-854-403

Query Match          61.4%; Score 387.5; DB 15; Length 93;
Best Local Similarity 74.8%; Pred. No. 2.1e-28;
Matches 80; Conservative 0; Mismatches 0; Indels 27; Gaps 1;

QY 14 LATVYVDVLKSGRDYVSQFEGSALGKQLNLKLLDNWDSVTSTFSKLRQLGPGVTQEPWD 73
Db 1 LATVYVDVLK-----DSVTSTFSKLRQLGPGVTQEPWD 33

QY 74 NLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWOEEMELYRQKVE 120
Db 34 NLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWOEEMELYRQKVE 80

RESULT 2
US-10-038-854-405
; Sequence 405, Application US/10038854
; Publication No. US20040022781A1
; GENERAL INFORMATION:
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Li, Li
; APPLICANT: Wolenc, Adam R
; APPLICANT: Vernet, Corine
; APPLICANT: Eisen, Andrew J
; APPLICANT: Liu, Xiaohong
; APPLICANT: Malvankar, Uriel M
; APPLICANT: Shimkets, Richard A
; APPLICANT: Tchernev, Vellizar
; APPLICANT: Spaderna, Steven K
; APPLICANT: Gorman, Linda
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Gangolli, Esha A
; APPLICANT: Guo, Xiaojia S
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Rastelli, Luca
; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc
; APPLICANT: Burgees, Catherine E
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gunther, Erik
; APPLICANT: Smithson, Glenda
; APPLICANT: Millet, Isabelle
```

```
; APPLICANT: MacDougall, John R
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-230
; CURRENT APPLICATION NUMBER: US/10/038,854
; CURRENT FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/258,928
; PRIOR FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: 60/259,415
; PRIOR FILING DATE: 2001-01-02
; PRIOR APPLICATION NUMBER: 60/259,785
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/269,814
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/279,832
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,833
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,863
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/283,889
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,447
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/286,583
; PRIOR FILING DATE: 2001-04-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 405
; LENGTH: 93
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-038-854-405

Query Match          61.4%; Score 387.5; DB 15; Length 93;
Best Local Similarity 74.8%; Pred. No. 2.1e-28;
Matches 80; Conservative 0; Mismatches 0; Indels 27; Gaps 1;

QY 14 LATVYVDVLKSGRDYVSQFEGSALGKQLNLKLLDNWDSVTSTFSKLRQLGPGVTQEPWD 73
Db 1 LATVYVDVLK-----DSVTSTFSKLRQLGPGVTQEPWD 33

QY 74 NLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWOEEMELYRQKVE 120
Db 34 NLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWOEEMELYRQKVE 80

RESULT 3
US-10-424-599-276814
; Sequence 276814, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 276814
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_91983C.1.pep
US-10-424-599-276814

Query Match          50.4%; Score 318; DB 15; Length 86;
Best Local Similarity 98.4%; Pred. No. 5.5e-22;
Matches 61; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```



```

RESULT 8
US-10-465-789A-50
; Sequence 50, Application US/10465789A
; Publication No. US20040053384A1
; GENERAL INFORMATION:
; APPLICANT: Sligar, Stephen G
; APPLICANT: Bayburt, Timothy H
; APPLICANT: Schuler, Mary A
; APPLICANT: Civjan, Natanya R
; APPLICANT: Yelena V. Grinkova
; APPLICANT: Ilia G. Denisov
; TITLE OF INVENTION: Membrane Scaffold Proteins
; FILE REFERENCE: 87-00A
; CURRENT APPLICATION NUMBER: US/10/465,789A
; CURRENT FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: 09/990,087
; PRIOR FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 60/252,233
; PRIOR FILING DATE: 2000-11-20
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 50
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: artificial sequence of Helix 2
US-10-465-789A-50

Query Match      18.7%; Score 118; DB 15; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.00039;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 66 PVTQEFWDNLEKETEGLRQEMS 87
|||||
DB 1 PVTQEFWDNLEKETEGLRQEMS 22

RESULT 9
US-10-465-789A-49
; Sequence 49, Application US/10465789A
; Publication No. US20040053384A1
; GENERAL INFORMATION:
; APPLICANT: Sligar, Stephen G
; APPLICANT: Bayburt, Timothy H
; APPLICANT: Schuler, Mary A
; APPLICANT: Civjan, Natanya R
; APPLICANT: Yelena V. Grinkova
; APPLICANT: Ilia G. Denisov
; TITLE OF INVENTION: Membrane Scaffold Proteins
; FILE REFERENCE: 87-00A
; CURRENT APPLICATION NUMBER: US/10/465,789A
; CURRENT FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: 09/990,087
; PRIOR FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 60/252,233
; PRIOR FILING DATE: 2000-11-20
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 49
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: artificial sequence of Helix 1
US-10-465-789A-49

Query Match      17.7%; Score 112; DB 15; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 LKLLDNWDSVTSTFSLKREQLG 65
|||||

Db 1 LKLLDNWDSVTSTFSLKREQLG 22

RESULT 10
US-10-487-096-2
; Sequence 2, Application US/10487096
; Publication No. US20040197823A1
; GENERAL INFORMATION:
; APPLICANT: GENFIT SA
; TITLE OF INVENTION: Compositions and Methods for the assay of AA4RP
; FILE REFERENCE: B0157W0
; CURRENT APPLICATION NUMBER: US/10/487,096
; CURRENT FILING DATE: 2004-02-19
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 2
; LENGTH: 95
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Description of artificial sequence : synthetic peptide AA4RP
US-10-487-096-2

Query Match      15.3%; Score 96.5; DB 17; Length 95;
Best Local Similarity 29.5%; Pred. No. 0.23;
Matches 23; Conservative 21; Mismatches 31; Indels 3; Gaps 2;

QY 24 DSGRDYVSQFEGSALGKQNLKLLDNWDSVTSTFSLKREQLGPTVQEFWDNLEKETEGLR 83
|||||
DB 18 DKGR--VQIHHQKQWRE-PATLKDSLEQDLNNMNFLEKLRPLSGSEAPRLPQDPVGWR 74
|||
QY 84 QEMSKDLEEVKAKVQPYL 101
|||
DB 75 RQLQEELEEVKARLQPYM 92

RESULT 11
US-10-142-238A-40
; Sequence 40, Application US/10142238A
; Publication No. US20030087819A1
; GENERAL INFORMATION:
; APPLICANT: Bielicki, John K.
; TITLE OF INVENTION: CYSTEINE-CONTAINING PEPTIDES HAVING OXIDANT PROPERTIES
; FILE REFERENCE: IB-1705
; CURRENT APPLICATION NUMBER: US/10/142,238A
; CURRENT FILING DATE: 2002-08-19
; PRIOR APPLICATION NUMBER: US 60/289,944
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 40
; LENGTH: 30
; TYPE: PRT
; ORGANISM: ARTIFICIAL SEQUENCE
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(30)
; OTHER INFORMATION: HUMAN GENETIC ORIGIN
US-10-142-238A-40

Query Match      14.9%; Score 94; DB 14; Length 30;
Best Local Similarity 95.0%; Pred. No. 0.096;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 42 LNLKLLDNWDSVTSTFSLKLR 61
|||||
DB 11 LPLKLLDNWDSVTSTFSLKLR 30

RESULT 12
US-10-142-238A-41
; Sequence 41, Application US/10142238A
; Publication No. US20030087819A1
```

```

; PRIOR APPLICATION NUMBER: PCT/GB02/02022
; PRIOR FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: GB0110790.3
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: GB0118385.4
; PRIOR FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: GB0119791.2
; PRIOR FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: GB0120045.0
; PRIOR FILING DATE: 2001-08-16
; PRIOR APPLICATION NUMBER: GB0128062.7
; PRIOR FILING DATE: 2001-11-22
; NUMBER OF SEQ ID NOS: 156
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 37
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-700-340-37

Query Match      14.1%; Score 89; DB 17; Length 16;
Best Local Similarity 100.0%; Pred.No. 0.13;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      62  EQLGPTQEFWDNLEK 77
      |||||
DB      1  EQLGPTQEFWDNLEK 16

RESULT 15
US-10-142-238A-39
; Sequence 39, Application US/10142238A
; Publication No. US20030087819A1
; GENERAL INFORMATION:
; APPLICANT: Bielski, John K.
; TITLE OF INVENTION: CYSTEINE-CONTAINING PEPTIDES HAVING OXIDANT PROPERTIES
; FILE REFERENCE: IB-1705
; CURRENT APPLICATION NUMBER: US/10/142,238A
; CURRENT FILING DATE: 2002-08-19
; PRIOR APPLICATION NUMBER: US 60/289,944
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 39
; LENGTH: 30
; TYPE: PRT
; ORGANISM: ARTIFICIAL SEQUENCE
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(18)
; OTHER INFORMATION: HUMAN GENETIC ORIGIN
US-10-142-238A-39

Query Match      14.1%; Score 89; DB 14; Length 30;
Best Local Similarity 90.0%; Pred.No. 0.28;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      42  LNLKLLDNWDSVTSTFSKLR 61
      |||
DB      11  LPLKLCNWDSDVTSTFSKLR 30

Search completed: December 21, 2004, 12:43:40
Job time : 143 secs

```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2004, 13:06:21 ; Search time 38 Seconds
(without alignments)
209.425 Million cell updates/sec

Title: US-09-803-918A-2_COPY_25_144

Perfect score: 631

Sequence: 1 DEPPQSPWRVKDLATVYVD.....LDDFQKQWEMELYRQKVE 120

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 327303

Minimum DB seq length: 0

Maximum DB seq length: 120

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/1/iaa/5A COMB.pep.*

2: /cgn2_6/ptodata/1/iaa/5B COMB.pep.*

3: /cgn2_6/ptodata/1/iaa/6A COMB.pep.*

4: /cgn2_6/ptodata/1/iaa/6B COMB.pep.*

5: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pep.*

6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	198	31.4	44	2	US-08-292-870-2
2	188	29.8	64	2	US-08-292-870-1
3	138	21.9	32	2	US-08-292-870-3
4	124	19.7	25	2	US-08-292-870-4
5	91	14.4	16	1	US-07-959-946-5
6	91	14.4	16	1	US-08-333-577-5
7	91	14.4	16	5	PCT-US92-08634-5
8	78	12.4	110	1	US-07-849-389-7
9	76	12.0	107	1	US-08-182-175A-105
10	76	12.0	107	1	US-08-474-633A-92
11	76	12.0	107	4	US-08-823-771-92
12	76	12.0	107	5	PCT-US92-06412-105
13	72	11.4	105	4	US-09-513-999C-6681
14	67.5	10.7	87	1	US-08-685-764-4
15	67	10.6	13	5	PCT-US94-01234-47
16	66.5	10.5	99	4	US-09-543-681A-8306
17	66.5	10.5	105	1	US-08-241-853-11
18	66.5	10.5	105	2	US-08-850-917-11
19	65.5	10.4	99	4	US-09-147-875A-16
20	65.5	10.4	102	4	US-09-513-999C-8125
21	65	10.3	103	4	US-09-134-000C-4548
22	64.5	10.2	63	2	US-08-891-837B-3
23	64.5	10.2	63	2	US-08-891-837B-4
24	64	10.1	99	4	US-09-270-767-44037
25	63.5	10.1	50	3	US-09-695-458-9
26	63	10.0	103	4	US-09-270-767-57235
27	63	10.0	106	4	US-09-621-976-4439

28	62.5	9.9	101	2	US-08-710-749-1	Sequence 1, Appli
29	62	9.8	65	2	US-08-867-087B-34	Sequence 34, Appli
30	62	9.8	105	2	US-08-743-200-2	Sequence 2, Appli
31	62	9.8	110	4	US-09-205-258-1080	Sequence 1080, Ap
32	61.5	9.7	98	2	US-08-479-078-7	Sequence 7, Appli
33	61.5	9.7	99	2	US-08-710-749-10	Sequence 10, Appli
34	61.5	9.7	99	4	US-09-147-875A-11	Sequence 11, Appli
35	61	9.7	66	2	US-08-867-087B-70	Sequence 70, Appli
36	61	9.7	100	4	US-09-147-875A-12	Sequence 12, Appli
37	60.5	9.6	37	3	US-09-695-458-15	Sequence 15, Appli
38	60	9.5	80	3	US-09-183-861-61	Sequence 61, Appli
39	60	9.5	80	3	US-09-022-765-61	Sequence 61, Appli
40	60	9.5	80	4	US-09-071-035-134	Sequence 134, App
41	60	9.5	80	4	US-09-551-974A-61	Sequence 61, Appli
42	60	9.5	80	4	US-09-565-501A-61	Sequence 61, Appli
43	60	9.5	80	4	US-09-639-206A-61	Sequence 61, Appli
44	60	9.5	80	4	US-09-874-923-61	Sequence 61, Appli
45	60	9.5	85	4	US-09-134-000C-3703	Sequence 3703, Ap

ALIGNMENTS

RESULT 1
US-08-292-870-2
; Sequence 2, Application US/08292870
; Patent No. 5814467
; GENERAL INFORMATION:
; APPLICANT: Curtiss, Linda K
; APPLICANT: Banka, Carole L
; APPLICANT: Bonnet, David J
; APPLICANT: Smith, Richard S
; TITLE OF INVENTION: APO AI POLYPEPTIDES, DIAGNOSTIC METHODS
; TITLE OF INVENTION: AND SYSTEMS FOR QUANTIFYING APO AI, AND THERAPEUTIC
; TITLE OF INVENTION: METHODS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of
; ADDRESSEE: Patent Counsel
; STREET: 10666 No. 5814467th Torrey Pines Road., TPC 8
; CITY: La Jolla
; STATE: California
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/292,870
; FILING DATE: 17-AUG-1994
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/534,761
; FILING DATE: 07-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/711,333
; FILING DATE: 06-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US 91/04038
; FILING DATE: 07-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: TSRI 210.1 D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 44 amino acids
; TYPE: amino acid

;
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-08-292-870-2

Query Match 31.4%; Score 198; DB 2; Length 44;
Best Local Similarity 100.0%; Pred. No. 1.2e-12; Indels 0; Gaps 0;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 RQMSKDLSEVKAKVQPYLDDDFOKKQWQEMELYRQKVE 120
Db 1 RQMSKDLSEVKAKVQPYLDDDFOKKQWQEMELYRQKVE 38

RESULT 2

US-08-292-870-1
; Sequence 1, Application US/08292870
; Patent No. 5814467
; GENERAL INFORMATION:
; APPLICANT: Curtiss, Linda K
; APPLICANT: Banka, Carole L
; APPLICANT: Bonnet, David J
; APPLICANT: Smith, Richard S
; TITLE OF INVENTION: APO AI POLYPEPTIDES, DIAGNOSTIC METHODS
; TITLE OF INVENTION: AND SYSTEMS FOR QUANTIFYING APO AI, AND THERAPEUTIC
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of
; ADDRESSEE: Patent Counsel
; STREET: 10666 No. 5814467th Torrey Pines Road., TPC 8
; CITY: La Jolla
; STATE: California
; COUNTRY: US
; ZIP: 92037

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/292,870
; FILING DATE: 17-AUG-1994
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/534,761
; FILING DATE: 07-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/711,333
; FILING DATE: 06-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US 91/04038
; FILING DATE: 07-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: TSRI 210.1 D1
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 1:
; FILING DATE: 07-JUN-1991
; SEQUENCE CHARACTERISTICS:
; TYPE: amino acid
; LENGTH: 64 amino acids
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal

US-08-292-870-1
Query Match 29.8%; Score 188; DB 2; Length 64;
Best Local Similarity 100.0%; Pred. No. 1.9e-11;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 85 EMSKDLSEVKAKVQPYLDDDFOKKQWQEMELYRQKVE 120
Db 1 EMSKDLSEVKAKVQPYLDDDFOKKQWQEMELYRQKVE 36

RESULT 3

US-08-292-870-3
; Sequence 3, Application US/08292870
; Patent No. 5814467
; GENERAL INFORMATION:
; APPLICANT: Curtiss, Linda K
; APPLICANT: Banka, Carole L
; APPLICANT: Bonnet, David J
; APPLICANT: Smith, Richard S
; TITLE OF INVENTION: APO AI POLYPEPTIDES, DIAGNOSTIC METHODS
; TITLE OF INVENTION: AND SYSTEMS FOR QUANTIFYING APO AI, AND THERAPEUTIC
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of
; ADDRESSEE: Patent Counsel
; STREET: 10666 No. 5814467th Torrey Pines Road., TPC 8
; CITY: La Jolla
; STATE: California
; COUNTRY: US
; ZIP: 92037

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/292,870
; FILING DATE: 17-AUG-1994
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/534,761
; FILING DATE: 07-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/711,333
; FILING DATE: 06-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US 91/04038
; FILING DATE: 07-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: TSRI 210.1 D1
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 3:
; FILING DATE: 07-JUN-1991
; SEQUENCE CHARACTERISTICS:
; TYPE: amino acid
; LENGTH: 32 amino acids
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Region
; LOCATION: 11
; OTHER INFORMATION: /note="Xaa can be either E (Glu)
; OTHER INFORMATION: or F (Phe)"

US-08-292-870-3

Query Match 21.9%; Score 138; DB 2; Length 32;
Best Local Similarity 96.3%; Pred. No. 6.1e-07;
Matches 26; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 94 KAKVQPYLDDDFOKKQWQEMELYRQKVE 120
Db 1 KAKVQPYLDDDFOKKQWQEMELYRQKVE 27


```
RESULT 4
US-08-292-870-4
; Sequence 4, Application US/08292870
; Patent No. 5814467
; GENERAL INFORMATION:
; APPLICANT: Curtiss, Linda K
; APPLICANT: Banks, Carole L
; APPLICANT: Bonnet, David J
; APPLICANT: Smith, Richard S
; TITLE OF INVENTION: APO AI POLYPEPTIDES, DIAGNOSTIC METHODS
; TITLE OF INVENTION: AND SYSTEMS FOR QUANTIFYING APO AI, AND THERAPEUTIC
; TITLE OF INVENTION: METHODS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of
; ADDRESSEE: Patent Counsel
; STREET: 10666 No. 581467th Torrey Pines Road., TPC 8
; CITY: La Jolla
; STATE: California
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/292,870
; FILING DATE: 17-AUG-1994
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/534,761
; FILING DATE: 07-JUN-1990
; PRIOR APPLICATION DATA: US 07/711,333
; FILING DATE: 06-JUN-1991
; PRIOR APPLICATION DATA: PCT/US 91/04038
; APPLICATION NUMBER: PCT/US 91/04038
; FILING DATE: 07-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: TSRI 210.1 D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-08-292-870-4
Query Match 19.7%; Score 124; DB 2; Length 25;
Best Local Similarity 96.0%; Pred. No. 1.1e-05;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 87 SKDLEEVKAKVQPYLDDFQKKWQEE 111
|||||
Db 1 SKDLEEVKAKVQPYLDDFQKKWQEE 25

RESULT 5
US-07-959-946-5
; Sequence 5, Application US/07959946
; Patent No. 5408038
; GENERAL INFORMATION:
; APPLICANT: Smith, Richard K.
; APPLICANT: Koduri, Raju
; APPLICANT: Young, Stephen G.
; TITLE OF INVENTION: Lipoprotein Assays Using Antibodies to a
; TITLE OF INVENTION: Pan Native Epitope and Recombinant Antigens
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dressler, Goldsmith, Shore, Sutker &
; ADDRESSEE: Milnamow, Ltd.
; STREET: 180 No. 5786206th Stetson, Suite 4700
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
; APPLICANT: Witztum, Joseph L.
; APPLICANT: Curtiss, Linda K.
; TITLE OF INVENTION: Lipoprotein Assays Using Antibodies to a
; TITLE OF INVENTION: Pan Native Epitope and Recombinant Antigens
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dressler, Goldsmith, Shore, Sutker &
; ADDRESSEE: Milnamow, Ltd.
; STREET: 180 No. 5408038th Stetson, Suite 4700
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/959,946
; FILING DATE: 19921008
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/901,706
; FILING DATE: 18-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Ganson, Edward P.
; REGISTRATION NUMBER: 29,381
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312)616-5400
; TELEFAX: (312)616-5460
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-959-946-5
Query Match 14.4%; Score 91; DB 1; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.01;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 96 KVQPYLDDFQKKWQEE 111
|||||
Db 1 KVQPYLDDFQKKWQEE 16

RESULT 6
US-08-333-577-5
; Sequence 5, Application US/08333577
; Patent No. 5786206
; GENERAL INFORMATION:
; APPLICANT: Smith, Richard K.
; APPLICANT: Koduri, Raju
; APPLICANT: Young, Stephen G.
; APPLICANT: Witztum, Joseph L.
; APPLICANT: Curtiss, Linda K.
; TITLE OF INVENTION: Lipoprotein Assays Using Antibodies to a
; TITLE OF INVENTION: Pan Native Epitope and Recombinant Antigens
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dressler, Goldsmith, Shore, Sutker &
; ADDRESSEE: Milnamow, Ltd.
; STREET: 180 No. 5786206th Stetson, Suite 4700
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
;
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/333,577
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Gamson, Edward P.
; REGISTRATION NUMBER: 29,381
; REFERENCE/DOCKET NUMBER: SCRF 234.0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312)616-5400
; TELEFAX: (312)616-5460
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-333-577-5

Query Match      14.4%; Score 91; DB 1; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.01;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      96 KVQPYLDDFQKKWQEE 111
DB      1 KVQPYLDDFQKKWQEE 16

;
; RESULT 7
; Sequence 5, Application PC/TUS9208634
; GENERAL INFORMATION:
; APPLICANT: Smith, Richard K.
; APPLICANT: Koduri, Raju
; APPLICANT: Young, Stephen G.
; APPLICANT: Witczum, Joseph L.
; APPLICANT: Curtiss, Linda K.
; TITLE OF INVENTION: Lipoprotein Assays Using Antibodies to a
; TITLE OF INVENTION: Pan Native Epitope and Recombinant Antigens
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dressler, Goldemith, Shore, Sutker &
; ADDRESSEE: Milnamow, Ltd.
; STREET: 180 North Stetson, Suite 4700
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/08634
; FILING DATE: 19921009
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION NUMBER: US 07/901,706
; FILING DATE: 18-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Gamson, Edward P.
; REGISTRATION NUMBER: 29,381
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312)616-5400
; TELEFAX: (312)616-5460
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: AMINO ACID
```

```
;
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US92-08634-5

Query Match      14.4%; Score 91; DB 5; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.01;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      96 KVQPYLDDFQKKWQEE 111
DB      1 KVQPYLDDFQKKWQEE 16

;
; RESULT 8
; US-07-849-389-7
; Sequence 7, Application US/07849389
; Patent No. 5525493
; GENERAL INFORMATION:
; APPLICANT: HORNES, Erik
; APPLICANT: UHLEN, Mathias
; TITLE OF INVENTION: CLONING METHOD AND KIT
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/849,389
; FILING DATE: 19920519
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16787/168/DFBC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 110 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-849-389-7

Query Match      12.4%; Score 78; DB 1; Length 110;
Best Local Similarity 29.0%; Pred. No. 2.1;
Matches 18; Conservative 13; Mismatches 31; Indels 0; Gaps 0;

QY      58 SKLREQLGPEVTQEFWDNLEKETGLRQEMSKDLSEVKAKVQPYLDDFQKKWQEEMLYRQ 117
DB      8 SELEQLTPVAETEARLSKELQAAEAFLGADMEDVGRGLVQYRGEVQAMLGQSTELRV 67

;
; QY      118 KV 119
; DB      68 RL 69

;
; RESULT 9
; US-08-182-175A-105
; Sequence 105, Application US/08182175A
; Patent No. 5559223
; GENERAL INFORMATION:
; APPLICANT: Saverio Carl Falco
```

```
/
/ APPLICANT: Sharon J. Keeler
/ APPLICANT: Janet A. Rice
/ TITLE OF INVENTION: Synthetic Storage Proteins with Defined Structure Containing B
/ NUMBER OF SEQUENCES: 113
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: E. I. du Pont de Nemours and Company
/ STREET: 1007 Market Street
/ CITY: Wilmington
/ STATE: Delaware
/ COUNTRY: USA
/ ZIP: 19898
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy Disk
/ COMPUTER: Macintosh
/ OPERATING SYSTEM: Macintosh System, 6.0
/ SOFTWARE: Microsoft Word, 4.0
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/182,175A
/ FILING DATE:
/ CLASSIFICATION: 800
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 07/743,006
/ FILING DATE: 9 August 1991
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Linda Axamethy Floyd
/ REGISTRATION NUMBER: 33,692
/ REFERENCE/DOCKET NUMBER: BB-1031
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (302) 992-4929
/ TELEFAX: (302) 892-7949
/ TELEX: 835420
/ INFORMATION FOR SEQ ID NO: 105:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 107 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/
/ US-08-182-175A-105
/
/ Query Match 12.0%; Score 76; DB 1; Length 107;
/ Best Local Similarity 25.7%; Pred. No. 3.2;
/ Matches 18; Conservative 22; Mismatches 22; Indels 8; Gaps 2;
/
/ QY 59 KLREQLGPTVQTFW-----DNLEKETEGRLQRMKDLBEVKA--KVQPYLDDFQKKWQE 110
/ Db 7 KLKEEMAKMKDEMVKLKEEMKKLEEKVMEEKMKLEEKMKAMEDKMKWLEBKMKKLE 66
/
/ QY 111 EMELYRQKVE 120
/ Db 67 KMKVMEERKM 76
/
/ RESULT 10
/ US-08-474-633A-92
/ Sequence 92, Application US/08474633A
/ Patent No. 5773691
/ GENERAL INFORMATION:
/ APPLICANT: E. I. DU PONT DE NEMOURS AND
/ APPLICANT: COMPANY
/ TITLE OF INVENTION: CHIMERIC GENES AND
/ TITLE OF INVENTION: METHODS FOR INCREASING
/ TITLE OF INVENTION: INCREASING THE LYSINE
/ TITLE OF INVENTION: AND THREONINE CONTENT
/ TITLE OF INVENTION: OF THE SEEDS OF PLANTS
/ NUMBER OF SEQUENCES: 107
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: E. I. DU PONT DE NEMOURS
/ STREET: 1007 MARKET STREET
/ CITY: WILMINGTON
/ STATE: DELAWARE
/ COUNTRY: U.S.A.
/ ZIP: 19898
/
/ APPLICANT: Sharon J. Keeler
/ APPLICANT: Janet A. Rice
/ TITLE OF INVENTION: Synthetic Storage Proteins with Defined Structure Containing B
/ NUMBER OF SEQUENCES: 113
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: E. I. du Pont de Nemours and Company
/ STREET: 1007 Market Street
/ CITY: Wilmington
/ STATE: Delaware
/ COUNTRY: USA
/ ZIP: 19898
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy Disk
/ COMPUTER: Macintosh
/ OPERATING SYSTEM: Macintosh System, 6.0
/ SOFTWARE: Microsoft Word, 4.0
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/182,175A
/ FILING DATE:
/ CLASSIFICATION: 800
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 07/743,006
/ FILING DATE: 9 August 1991
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Linda Axamethy Floyd
/ REGISTRATION NUMBER: 33,692
/ REFERENCE/DOCKET NUMBER: BB-1031
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (302) 992-4929
/ TELEFAX: (302) 892-7949
/ TELEX: 835420
/ INFORMATION FOR SEQ ID NO: 105:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 107 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/
/ US-08-182-175A-105
/
/ Query Match 12.0%; Score 76; DB 1; Length 107;
/ Best Local Similarity 25.7%; Pred. No. 3.2;
/ Matches 18; Conservative 22; Mismatches 22; Indels 8; Gaps 2;
/
/ QY 59 KLREQLGPTVQTFW-----DNLEKETEGRLQRMKDLBEVKA--KVQPYLDDFQKKWQE 110
/ Db 7 KLKEEMAKMKDEMVKLKEEMKKLEEKVMEEKMKLEEKMKAMEDKMKWLEBKMKKLE 66
/
/ QY 111 EMELYRQKVE 120
/ Db 67 KMKVMEERKM 76
/
/ RESULT 11
/ US-08-823-771-92
/ Sequence 92, Application US/08823771
/ Patent No. 6459019
/ GENERAL INFORMATION:
/ APPLICANT: E. I. DU PONT DE NEMOURS AND
/ APPLICANT: COMPANY
/ TITLE OF INVENTION: CHIMERIC GENES AND
/ TITLE OF INVENTION: METHODS FOR INCREASING
/ TITLE OF INVENTION: INCREASING THE LYSINE
/ TITLE OF INVENTION: AND THREONINE CONTENT
/ NUMBER OF SEQUENCES: 107
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: E. I. DU PONT DE NEMOURS
/ STREET: 1007 MARKET STREET
/ CITY: WILMINGTON
/ STATE: DELAWARE
/ COUNTRY: U.S.A.
/ ZIP: 19898
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy Disk
/ COMPUTER: IBM PC COMPATIBLE
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: MICROSOFT WORD VERSION 2.0C
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/823,771
/ FILING DATE: 24-Mar-1997
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/474,633
/ FILING DATE: <Unknown>
/ ATTORNEY/AGENT INFORMATION:
/ NAME: BARBARA C. SIEGELL
/ REGISTRATION NUMBER: 30,684
/
```

```
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: FLOPPY DISK
/ COMPUTER: IBM PC COMPATIBLE
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: MICROSOFT WORD VERSION 2.0C
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/474,633A
/ FILING DATE:
/ CLASSIFICATION: 800
/ ATTORNEY/AGENT INFORMATION:
/ NAME: BARBARA C. SIEGELL
/ REGISTRATION NUMBER: 30,684
/ REFERENCE/DOCKET NUMBER: BB-1037-C
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 302-992-4931
/ TELEFAX: 302-773-0164
/ TELEX: 835420
/ INFORMATION FOR SEQ ID NO: 92:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 107 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/
/ US-08-474-633A-92
/
/ Query Match 12.0%; Score 76; DB 1; Length 107;
/ Best Local Similarity 25.7%; Pred. No. 3.2;
/ Matches 18; Conservative 22; Mismatches 22; Indels 8; Gaps 2;
/
/ QY 59 KLREQLGPTVQTFW-----DNLEKETEGRLQRMKDLBEVKA--KVQPYLDDFQKKWQE 110
/ Db 7 KLKEEMAKMKDEMVKLKEEMKKLEEKVMEEKMKLEEKMKAMEDKMKWLEBKMKKLE 66
/
/ QY 111 EMELYRQKVE 120
/ Db 67 KMKVMEERKM 76
/
/ RESULT 11
/ US-08-823-771-92
/ Sequence 92, Application US/08823771
/ Patent No. 6459019
/ GENERAL INFORMATION:
/ APPLICANT: E. I. DU PONT DE NEMOURS AND
/ APPLICANT: COMPANY
/ TITLE OF INVENTION: CHIMERIC GENES AND
/ TITLE OF INVENTION: METHODS FOR INCREASING
/ TITLE OF INVENTION: INCREASING THE LYSINE
/ TITLE OF INVENTION: AND THREONINE CONTENT
/ NUMBER OF SEQUENCES: 107
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: E. I. DU PONT DE NEMOURS
/ STREET: 1007 MARKET STREET
/ CITY: WILMINGTON
/ STATE: DELAWARE
/ COUNTRY: U.S.A.
/ ZIP: 19898
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: FLOPPY DISK
/ COMPUTER: IBM PC COMPATIBLE
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: MICROSOFT WORD VERSION 2.0C
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/823,771
/ FILING DATE: 24-Mar-1997
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/474,633
/ FILING DATE: <Unknown>
/ ATTORNEY/AGENT INFORMATION:
/ NAME: BARBARA C. SIEGELL
/ REGISTRATION NUMBER: 30,684
/
```

```

; REFERENCE/DOCKET NUMBER: BB-1037-C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-992-4931
; TELEFAX: 302-773-0164
; TELEX: 835420
; INFORMATION FOR SEQ ID NO: 92:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 107 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 92:
US-08-823-771-92

Query Match 12.0%; Score 76; DB 4; Length 107;
Best Local Similarity 25.7%; Pred. No. 3.2;
Matches 18; Conservative 22; Mismatches 22; Indels 8; Gaps 2;

QY 59 KLRQLGPGVTQFW-----DNLEKETEGLRQEMSKDLSEVKA--KVQPYLDDFOKKWQE 110
DB 7 KLKEEMAKMKDEMVKLKEEMKLEEKMKVMEKMKLEEKMKAMEDKMKWLEEKMKKLEE 66

QY 111 EMELYRQKVE 120
DB 67 KMKVMEKMK 76

RESULT 12
PCT-US92-06412-105
; Sequence 105, Application PC/TUS9206412
; GENERAL INFORMATION:
; APPLICANT: Saverio Carl Falco
; APPLICANT: Sharon J. Keeler
; APPLICANT: Janet A. Rice
; TITLE OF INVENTION: Synthetic Storage Proteins with Defined Structure Containing B
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E.I. du Pont de Nemours and Company
; STREET: 1007 Market Street
; CITY: Wilmington
; STATE: Delaware
; COUNTRY: USA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: Macintosh System, 6.0
; SOFTWARE: Microsoft Word, 4.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/06412
; FILING DATE: 19920807
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/743,006
; FILING DATE: 9 August 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Linda Axamethy Floyd
; REGISTRATION NUMBER: 33,692
; REFERENCE/DOCKET NUMBER: BB-1031
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (302) 992-4929
; TELEFAX: (302) 892-7949
; TELEX: 835420
; INFORMATION FOR SEQ ID NO: 105:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 107 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US92-06412-105

Query Match 12.0%; Score 76; DB 5; Length 107;
Best Local Similarity 25.7%; Pred. No. 3.2;

; REFERENCE/DOCKET NUMBER: BB-1037-C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-992-4931
; TELEFAX: 302-773-0164
; TELEX: 835420
; INFORMATION FOR SEQ ID NO: 92:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 107 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 92:
US-09-513-999C-6681
; Sequence 6681, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513.999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 6681
; LENGTH: 105
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-513-999C-6681

Query Match 11.4%; Score 72; DB 4; Length 105;
Best Local Similarity 24.2%; Pred. No. 7.7;
Matches 23; Conservative 22; Mismatches 36; Indels 14; Gaps 3;

QY 39 GKQLNKLKLD-NWDSVTSFTSKLRQLGPGVTQFWDNLEKETEG---LRQEMSKDLSEVYK 94
DB 7 GKEQQLDIMNKYQQLSRSLDEILSRIAKETEIKDLSEQLTEGQIAAANEALKKDLGVY 66

QY 95 AKVQPYLDDFQ-----KKQEEEMELYRQKV 119
DB 67 SGLQBYLTGTIRGQATQQAQNECKRKLURDEKETLLQL 101

RESULT 14
US-08-685-764-4
; Sequence 4, Application US/08685764
; Patent No. 5800982
; GENERAL INFORMATION:
; APPLICANT: HASEGAWA, AKIRA
; APPLICANT: MAKI, NOBORU
; APPLICANT: YAGI, SHINTARO
; APPLICANT: KASHIWAKUMA, TOMIKO
; TITLE OF INVENTION: ANTIGENIC PEPTIDES FOR GROUPING
; TITLE OF INVENTION: HEPATITIS C VIRUS, KIT COMPRISING THE SAME AND
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
; ADDRESSEE: CUSHMAN
; STREET: 130 WATER STREET
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
```

;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/685,764
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/092,192
; FILING DATE: 15-JUL-1993
; APPLICATION NUMBER: JP 212061/92
; FILING DATE: 16-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 316634/92
; FILING DATE: 30-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 316635/92
; FILING DATE: 30-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 104754/93
; FILING DATE: 30-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: BUCKLEY, LINDA M.
; REGISTRATION NUMBER: 31003
; REFERENCE/DOCKET NUMBER: 42822
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 523-3400
; TELEFAX:
; TELEX: 200291 STRE UR
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 87 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-685-764-4

Query Match 10.7%; Score 67.5; DB 1; Length 87;
Best Local Similarity 24.2%; Pred. No. 17; Mismatches 39; Indels 11; Gaps 3;
Matches 23; Conservative 22;
QY 13 DLATVYVDVLKDSGRDVVSQFEGSALGKQLNLKLDNWDSTVTFKRLRQLGPTQEFW 72
Db 1 EFATGCVSII--GRLLHQRAVAPDKVLYEAFDEMECASRAALIEB-----CQRTA 52
QY 73 DNLKETEGRLRQEMSKDLBVKAKVQ---PYLDDF 104
Db 53 EMLKSKIQGLLQQAQKQAQDIKPAVQTSPKVKQF 87

RESULT 15
PCT-US94-01234-47
; Sequence 47, Application PC/TUS9401234
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: METHODS FOR PRODUCING POLYPEPTIDE
; TITLE OF INVENTION: BINDING SITES
; NUMBER OF SEQUENCES: 76
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/01234
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/084,542
; FILING DATE: 28-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/012,566
; FILING DATE: 02-FEB-1993
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

;
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; PCT-US94-01234-47
Query Match 10.6%; Score 67; DB 5; Length 13;
Best Local Similarity 92.3%; Pred. No. 1.8;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 99 PYLDDFQKKWQEE 111
Db 1 PYLDDFQKKWQEE 13
Search completed: December 21, 2004, 13:19:49
Job time : 39 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2004, 12:24:55 ; Search time 191 Seconds
(without alignments)
328.355 Million cell updates/sec

Title: US-09-803-918A-2_COPY_25_133
Perfect score: 576
Sequence: 1 DEPPQSPWDRVKOLATVYVD.....LEEYKAKVQPLDFFQKKW 109

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 1825181 seqs, 575374646 residues
Total number of hits satisfying chosen parameters: 337476

Minimum DB seq length: 0
Maximum DB seq length: 109

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot_02: *
1: uniprot_sprot: *
2: uniprot_trembl: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	247	42.9	56	2	O02762 ovis aries
2	217	37.7	82	2	Q29248 sus scrofa
3	183	31.8	79	2	O6LD50 mus sp. apo
4	183	31.8	79	2	AAB35539 mus sp. a
5	141	24.5	26	2	Q9UCT8
6	106	18.4	34	2	Q9PRR6
7	101	17.5	20	1	APAI_ERYPA
8	101	17.5	30	2	Q9QV04
9	83	14.4	107	1	APE_SAI5C
10	82	14.2	107	1	APE_MACMU
11	77.5	13.5	96	2	Q7P280
12	74.5	12.9	102	2	Q9PXP4
13	73	12.7	76	2	O6WHU6
14	73	12.7	76	2	AAQ64377
15	73	12.7	93	2	O81R54
16	73	12.7	93	2	AAT31321
17	72	12.5	102	2	Q29258
18	68	11.8	90	2	Q9LCR0
19	68	11.8	101	2	O6BMU6
20	67	11.6	95	2	P89597
21	66.5	11.5	79	2	Q9CHE9
22	66	11.5	80	2	O6NK47
23	66	11.5	80	2	CAE48695
24	66	11.5	104	2	Q54892
25	65.5	11.4	95	2	O82VG7
26	65	11.3	102	2	O86F03
27	64	11.1	95	2	P89581
28	63.5	11.0	79	2	O8FXA5
29	63.5	11.0	96	2	Q9T5S1
30	63.5	11.0	105	2	O85812
31	63	10.9	69	1	SEMI_ASHGO

32	63	10.9	69	2	AAS52934	Aas52934 ashbya go
33	63	10.9	77	2	Q93K19	Q93K19 lactococcus
34	63	10.9	99	2	Q9PS33	Q9PS33 gallus gall
35	63	10.9	107	1	YC51_HAEIN	O57089 haemophilus
36	62.5	10.9	78	2	Q9RH71	Q9RH71 streptococc
37	62.5	10.9	92	2	Q73LU4	Q73LU4 treponema d
38	62.5	10.9	92	2	AAS12283	Aas12283 treponema
39	62.5	10.9	96	2	Q97AV3	Q97AV3 thermoplasma
40	62.5	10.9	98	2	Q8AWE6	Q8AWE6 oncorhynchus
41	62	10.8	92	2	O8TVR5	O8TVR5 methanopyru
42	62	10.8	95	2	P89588	P89588 human immun
43	62	10.8	95	2	P89596	P89596 human immun
44	62	10.8	96	2	Q7ZPL6	Q7ZPL6 human immun
45	61.5	10.7	75	2	Q6U9J3	Q6U9J3 bacteriophag

ALIGNMENTS

RESULT 1						
O02762	PRELIMINARY;	PRT;	56 AA.			
AC	O02762;					
DT	01-JUL-1997 (TREMBLrel. 04, Created)					
DT	01-JUL-1997 (TREMBLrel. 04, Last sequence update)					
DE	01-MAR-2004 (TREMBLrel. 26, Last annotation update)					
DE	Apolipoprotein A1 (Fragment)					
OS	Ovis aries (Sheep)					
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;					
OC	Caprinae; Ovis.					
OX	NCBI_TaxID:9940;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RA	MEDLINE=99098454; PubMed=9883985;					
RT	Robertson J.A., Bhattacharyya S., Ing N.H.;					
RT	"Tamoxifen up-regulates oestrogen receptor-alpha, c-fos and					
RT	glyceroldehyde 3-phosphate-dehydrogenase mRNAs in ovine endometrium.";					
RL	J. Steroid Biochem. Mol. Biol. 67:285-292(1998).					
DR	EMBL: U94720; AAB57840.1;					
DR	GO; GO:0005576; C:extracellular; IEA.					
DR	GO; GO:0008289; F:lipid binding; IEA.					
DR	GO; GO:0006869; P:lipid transport; IEA.					
DR	GO; GO:0042157; P:lipoprotein metabolism; IEA.					
DR	InterPro: IPR000074; Apolipoprotein.					
DR	InterPro: IPR009074; Apolipo_A_E_C3.					
DR	Pfam: PF01442; Apolipoprotein; 1.					
KW	Lipoprotein.					
FT	NON_TER	1	56			
FT	NON_TER	56	56			
SQ	SEQUENCE	56 AA;	6617 MW;	2AB38E08F1E8F1BC	CRC64;	

Query Match 42.9%; Score 247; DB 2; Length 56;
Best Local Similarity 82.1%; Pred. No. 9.5e-14;
Matches 46; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY	47	LDNWDSTVTSFKLRQLGPVTOEFWDNLEKEGLRQEMSKDL EEYKAKVQPYLD 102				
DB	1	LDNWDSTVTSFKLRQLGPVTOEFWDNLEKEGLRQEMSKDL EEYKAKVQPYLD 56				

RESULT 2

Q29248	PRELIMINARY;	PRT;	82 AA.			
ID	Q29248					
AC	Q29248;					
DT	01-NOV-1996 (TREMBLrel. 01, Created)					
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)					
DE	01-MAR-2004 (TREMBLrel. 26, Last annotation update)					
DE	Apolipoprotein A-I (Fragment)					
OS	Sus scrofa (Pig)					
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.					
OX	NCBI_TaxID:9823;					

```
RN SEQUENCE FROM N.A.
RP TISSUE=Small intestine;
RC MEDLINE=96327607; PubMed=8672129;
RA Winteroe A.K., Fredholm M., Davies W.;
RT "Evaluation and characterization of a porcine small intestine cDNA
RL library.";
RL Mamm. Genome 7:509-517(1996).
DR EMBL; F14858; CAA23298.1; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0008289; F:lipid binding; IEA.
DR GO; GO:0006869; P:lipid transport; IEA.
DR GO; GO:0042157; P:lipoprotein metabolism; IEA.
DR InterPro; IPR000074; Apolipoprotein.
DR InterPro; IPR009074; Apolipo_A_E_C3.
DR Pfam; PF01442; Apolipoprotein; I.
KW Lipoprotein.
FT NON_TER 82
SQ SEQUENCE 82 AA; 9168 MW; 24625C65CBFFEDD8 CRC64;

Query Match 37.7%; Score 217; DB 2; Length 82;
Best Local Similarity 70.7%; Pred. No. 5.2e-11;
Matches 41; Conservative 5; Mismatches 12; Indels 0; Gaps 0;

QY 2 BPPQSPWDRVKDLATVYVDVKDSGRDYVSQFSGSALGKLNKLLDNWDSVTSTFSK 59
Db 25 DDPXSPWDRVKDFTVTYVDAIKSGRDYXQAFASALGRHLNKLKLNKXDSLXTFTX 82

RESULT 3
Q6LD50 PRELIMINARY; PRT; 79 AA.
AC Q6LD50;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Apolipoprotein A1 homolog protein (Fragment).
GN Name=apolipoprotein A1 homolog;
OS Mus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10095;
RN [1]
RP MEDLINE=96062440; PubMed=7488287;
RA Merrill J.T., Rivkin E., Shen C., Lahita R.G.;
RT "Selection of a gene for apolipoprotein A1 using autoantibodies from a
RL patient with systemic lupus erythematosus.";
RL Arthritis Rheum. 38:1655-1659(1995).
DR EMBL; S80442; AAB35539.1; -.
DR InterPro; IPR000074; Apolipoprotein.
DR InterPro; IPR009074; Apolipo_A_E_C3.
KW Lipoprotein.
FT NON_TER 79
SQ SEQUENCE 79 AA; 9583 MW; 33CA72DA854A150A CRC64;

Query Match 31.8%; Score 183; DB 2; Length 79;
Best Local Similarity 80.0%; Pred. No. 3.8e-08;
Matches 32; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 70 EFWDNLEKETEGRLQEMSKDLSEVKAKVQPYLDDFOKKWQ 109
Db 1 DFWDNLEKETDWRQEMNKDLSEVKQKQPYLDEFQKKWK 40

RESULT 4
AAB35539 PRELIMINARY; PRT; 79 AA.
AC AAB35539;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE Apolipoprotein A1 homolog (Fragment).
OS Mus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10095;
RN [1]
RP MEDLINE=96062440; PubMed=7488287;
RA Merrill J.T., Rivkin E., Shen C., Lahita R.G.;
RT "Selection of a gene for apolipoprotein A1 using autoantibodies from a
RL patient with systemic lupus erythematosus.";
RL Arthritis Rheum. 38:1655-1659(1995).
DR EMBL; S80442; AAB35539.1; -.
DR InterPro; IPR000074; Apolipoprotein.
DR InterPro; IPR009074; Apolipo_A_E_C3.
KW Lipoprotein.
FT NON_TER 79
SQ SEQUENCE 79 AA; 9583 MW; 33CA72DA854A150A CRC64;

Query Match 31.8%; Score 183; DB 2; Length 79;
Best Local Similarity 80.0%; Pred. No. 3.8e-08;
Matches 32; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 70 EFWDNLEKETEGRLQEMSKDLSEVKAKVQPYLDDFOKKWQ 109
Db 1 DFWDNLEKETDWRQEMNKDLSEVKQKQPYLDEFQKKWK 40

RESULT 5
Q9UCT8 PRELIMINARY; PRT; 26 AA.
AC Q9UCT8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Sperm activating protein subunit I, apolipoprotein A1, SPAP subunit I
DE (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE=91369902; PubMed=1909888;
RA Akerlof E., Jorvall H., Slotte H., Pousette A.;
RT "Identification of apolipoprotein A1 and immunoglobulin as components
RT of a serum complex that mediates activation of human sperm motility.";
RL Biochemistry 30:8986-8990(1991).
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0008289; F:lipid binding; IEA.
DR GO; GO:0008869; P:lipid transport; IEA.
DR GO; GO:0042157; P:lipoprotein metabolism; IEA.
DR InterPro; IPR000074; Apolipoprotein.
DR InterPro; IPR009074; Apolipo_A_E_C3.
DR Pfam; PF01442; Apolipoprotein; I.
FT NON_TER 26
FT NON_TER 26
SQ SEQUENCE 26 AA; 2930 MW; 64921A333E768D27 CRC64;

Query Match 24.5%; Score 141; DB 2; Length 26;
Best Local Similarity 100.0%; Pred. No. 3.9e-05;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DEPPQSPWDRVKDLATVYVDVKDSG 26
Db 1 DEPPQSPWDRVKDLATVYVDVKDSG 26

RESULT 6
Q9PRR6 PRELIMINARY; PRT; 34 AA.
AC Q9PRR6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Apolipoprotein A1 (Fragment).
OS Anser anser (domestic goose).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```



```

OC Archosauria; Aves; Neognathae; Anseriformes; Anatidae; Anser.
RN NCBI_TaxID=8843;
RN SEQUENCE.
RX MEDLINE=96128192; PubMed=8536707;
RA Hermier D., Sellier N., Rousselot-Pailley D., Forgez P.;
RT "Characterization of apolipoproteins B-100, AI and C from plasma
RT lipoprotein in the goose, Anser anser. Evidence for a genetic
RT polymorphism in ApoC-like apolipoproteins."
RL Eur. J. Biochem. 234:586-591(1995).
DR PIR: S67972; S67972.
DR GO: GO:0005576; C:extracellular; IEA.
DR GO: GO:0008289; F:lipid binding; IEA.
DR GO: GO:0006869; P:lipid transport; IEA.
DR GO: GO:0042157; P:lipoprotein metabolism; IEA.
DR InterPro: IPR000074; Apolipoprotein.
DR Pfam: PF01442; Apolipoprotein_1.
DR NCBI_TaxID=9538;
SQ SEQUENCE 34 AA; 3838 MW; 4BE7DFA02BF1DE91 CRC64;

Query Match 18.4%; Score 106; DB 2; Length 34;
Best Local Similarity 54.5%; Pred. No. 0.05;
Matches 18; Conservative 9; Mismatches 6; Indels 0; Gaps 0;

Qy 2 EPQSPWDRVKDLATVYVDVLDKSGRDYVSQFE 34
: ||| ||| ||| ||| : ||| ||| |||
Db 1 DEQAPLDRLKDLVDVYLETVKASGKDYLAQFE 33

RESULT 7
APAL_BRYPA STANDARD; PRT; 20 AA.
AC P18647;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Apolipoprotein A-I (Apo-AI) (Fragment).
GN Name=APOA1;
OS Erythrocybus patas (Red guenon) (Cercopithecus patas).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC Cercopithecinae; Erythrocybus.
OX NCBI_TaxID=9538;
RN [1]
RN SEQUENCE.
RX MEDLINE=76184721; PubMed=178359;
RA Mahley R.W., Weisgraber K.H., Innerarity T., Brewer H.B. Jr.;
RT "Characterization of the plasma lipoproteins and apoproteins of the
RT Erythrocybus patas monkey."
RL Biochemistry 15:1928-1933(1976).
CC -|- FUNCTION: Participates in the reverse transport of cholesterol
CC from tissues to the liver for excretion by promoting cholesterol
CC efflux from tissues and by acting as a cofactor for the lecithin
CC cholesterol acyltransferase (LCAT).
CC -|- SUBCELLULAR LOCATION: Secreted.
CC -|- TISSUE SPECIFICITY: Major protein of plasma HDL, also found in
CC chylomicrons.
CC -|- SIMILARITY: Belongs to the apolipoprotein A1/A4/E family.
DR PIR: A05313; A05313.
DR InterPro: IPR009074; Apolipo_A_E_C3.
KW Cholesterol metabolism; Direct protein sequencing; HDL;
KW Lipid transport; Plasma.
FT NON TER 20
SQ SEQUENCE 20 AA; 2387 MW; 9C970997C7FC976A CRC64;

Query Match 17.5%; Score 101; DB 1; Length 20;
Best Local Similarity 85.0%; Pred. No. 0.072;
Matches 17; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DEPPQSPWDRVKDLATVYVD 20
||||| ||| ||| |||
Db 1 DEPPQTPWDRVKDLVTVYVE 20

us-09-803-918a-2_copy_25_133.sl.rup

RESULT 8
Q9QV04 PRELIMINARY; PRT; 30 AA.
AC Q9QV04;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Apolipoprotein E (Fragment).
OS Rattus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10118;
RN [1]
RN SEQUENCE.
RX MEDLINE=95355312; PubMed=7629028;
RA Motojima K., Goto S.;
RL J. Biochem. 117:597-602(1995).
DR GO: GO:0005576; C:extracellular; IEA.
DR GO: GO:0008289; F:lipid binding; IEA.
DR GO: GO:0006869; P:lipid transport; IEA.
DR GO: GO:0042157; P:lipoprotein metabolism; IEA.
DR InterPro: IPR000074; Apolipoprotein.
DR Pfam: PF01442; Apolipoprotein_1.
DR NCBI_TaxID=10118;
SQ SEQUENCE 30 AA; 3355 MW; 4A0103DCB6242BF3 CRC64;

Query Match 17.5%; Score 101; DB 2; Length 30;
Best Local Similarity 67.9%; Pred. No. 0.11;
Matches 19; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

Qy 4 PQSPWDRVKDLATVYVDVLDKSGRDYVS 31
||||| ||| ||| ||| : ||| ||| |||
Db 3 PQSQWDXVXDFATVYVDVLDKSGDXYS 30

RESULT 9
APE_SAISC STANDARD; PRT; 107 AA.
ID APE_SAISC
AC Q28595;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Apolipoprotein E (Apo-E) (Fragment).
GN Name=APOE;
OS Saimiri sciureus (Common squirrel monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Saimiri.
OX NCBI_TaxID=9521;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=96225955; PubMed=8635577;
RA Morelli L., Wei L., Amorim A., McDermid J., Abee C.R., Frangione B.,
RA Walker L.C., Levy E.;
RT "Cerebrovascular amyloidosis in squirrel monkeys and rhesus monkeys:
RT apolipoprotein E genotype."
RL PNAS Lett. 379:132-134(1996).
CC -|- FUNCTION: Mediates the binding, internalization, and catabolism of
CC lipoprotein particles. It can serve as a ligand for the LDL (apo
CC B/E) receptor and for the specific apo-E receptor (chylomicron
CC remnant) of hepatic tissues.
CC -|- SUBCELLULAR LOCATION: Secreted.
CC -|- TISSUE SPECIFICITY: Secreted in plasma.
CC -|- SIMILARITY: Belongs to the apolipoprotein A1/A4/E family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U52029; AAC50442.1; -.

```

FT	DOMAIN	89	99		LDL receptor binding (Potential).				
FT	DOMAIN	93	96		Heparin-binding (By similarity).				
FT	DOMAIN	11	>107		8 X 22 AA approximate tandem repeats.				
FT	REPEAT	11	32	1.					
FT	REPEAT	33	54	2.					
FT	REPEAT	55	76	3.					
FT	REPEAT	77	98	4.					
FT	REPEAT	99	>107	5.					
FT	NON TER	107	107						
SQ	SEQUENCE	107 AA;	12382 MW;	E1D38C32F5AACB23 CRC64;					
Query Match									
Best Local Similarity 14.2%; Score 82; DB 1; Length 107;									
Matches 16; Conservative 12; Mismatches 20; Indels 0; Gaps 0;									
Qy	58 SKLRQLGPTVQTQFNLEKETEGLRQEMSKDL EEVKAKVPQYLD D F Q K	105							
Dd	25 SELEQLSPVAETRLRSKELQAARIGADMEDVRSRLVQRSEVQ	72							
RESULT 11									
Q7P280	PRELIMINARY; PRT; 96 AA.								
ID	Q7P280								
AC	Q7P280;								
DT	01-MAR-2004 (TrEMBLrel. 26, Created)								
DT	01-MAR-2004 (TrEMBLrel. 26, Last sequence update)								
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)								
DE	Hypothetical cytosolic protein.								
GN	Names:FNU0007;								
OS	Fusobacterium nucleatum subsp. vincentii ATCC 49256.								
OC	Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;								
OC	Fusobacterium.								
OX	NCBI_TaxID=209882;								
RN	[1]								
RP	SEQUENCE FROM N.A.								
RC	STRAN=ATCC 49256;								
RA	Karpatrial V., Ivanova N., Anderson I., Reznik G., Bhattacharyya A.,								
RA	Gardner W.L., Mikhailova N., Larsen N., D'Souza M., Walunas T.,								
RA	Haselkorn R., Overbeek R., Kyriades N.;								
RL	Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.								
CC	-1- CAUTION: The sequence shown here is derived from an								
CC	EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is								
CC	preliminary data.								
DR	EMBL; AABF0100020; EAA23153.1; --								
DR	InterPro; IPR004238; LEA.								
DR	Pfam; PF02987; LEA 4; 1.								
KW	Hypothetical protein.								
SQ	SEQUENCE 96 AA; 10436 MW; 7568787BAFE86350 CRC64;								
Query Match									
Best Local Similarity 13.5%; Score 77.5; DB 2; Length 96;									
Matches 23; Conservative 20; Mismatches 47; Indels 5; Gaps 1;									
Qy	13 DLATVTVDVLKDSGRDVYGFSGALGKQLNLKLNDWNDSVTTFSKLRQLGPTVQTQFW	72							
Dd	5 DDLTGLILDEVKDFVADKAKELKDEAVTKVGELK-----DKATEKAGELKDKNVWDKAKELK	59							
Qy	73 DNLEKETEGLRQEMSKDL EEVKAKVPQYLD D F Q K	107							
Dd	60 EGAEKSTSELKDAEKAEEKDKISGADSLINK	94							
RESULT 12									
Q9XP4	PRELIMINARY; PRT; 102 AA.								
ID	Q9XP4								
AC	Q9XP4;								
DT	01-MAY-2000 (TrEMBLrel. 13, Created)								
DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)								
DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)								

```

OC Hepacivirus.
OX NCBI_TaxID=111103;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94245087; PubMed=7514558;
RA Tanaka T., Tsukiyama-Kohara K., Yamaguchi K., Yagi S., Tanaka S.,
RA Hasegawa A., Ohta Y., Hattori N., Kohara M.;
RT "Significance of a specific antibody assay for genotyping of hepatitis C
RT virus."
RL Hepatology 19:1347-1353(1994).
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
FT NON TER 102
SQ SEQUENCE 102 AA; 11389 MW; 267CA4C1F2F7F44E CRC64;

Query Match 12.9%; Score 74.5; DB 2; Length 102;
Best Local Similarity 25.3%; Pred. No. 82;
Matches 25; Conservative 22; Mismatches 41; Indels 11; Gaps 3;

QY 9 DRVKDLATVVVLKDSGRDYVSQFEGSALGKQLNLKLLDNWDSVTSTFSKLRQLGPT 68
DB 12 DRDPEFATGCVSII---GRHINQRAVAPDKVELYEAFDEMEECASRAALIEE-----G 63

QY 69 QEFWDNLEKETEGLRQEMSKDLEEVKAKVQ---PYLDDF 104
DB 64 QRIAEMLKSKIQLGLQQAASKQADIKPAVQTSWPKVEQF 102

RESULT 13
Q6WHJ6
ID Q6WHJ6 PRELIMINARY; PRT; 76 AA.
AC Q6WHJ6;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Hypothetical protein.
GN ORFNames=KVP40.0307;
OS Bacteriophage KVP40.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae.
OX NCBI_TaxID=75320;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2803260; PubMed=12923095;
RA Miller E., Heidelberg J., Eise J., Nelson W., Durkin A., Ciecko A.,
RA Feldblyum T., White O., Paulsen I., Nierman W., Lee J.,
RA Szczypinski B., Fraser C.;
RT "Complete genome sequence of the broad-host-range vibriophage KVP40:
RT comparative genomics of a T4-related bacteriophage.";
RL J. Bacteriol. 185:5220-5233(2003).
RN [2]
RP SEQUENCE FROM N.A.
RA Miller E., Lee J., Szczypinski B.;
RL Submitted (Apr-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY283928; AAQ64377.1; -.
KW Hypothetical protein.
SQ SEQUENCE 76 AA; 9023 MW; 873E2003C17C60C6 CRC64;

Query Match 12.7%; Score 73; DB 2; Length 76;
Best Local Similarity 24.3%; Pred. No. 79;
Matches 18; Conservative 12; Mismatches 24; Indels 20; Gaps 2;

QY 34 EGSALGKQLNLKLLDNWDSVTSTFSKLRQLGPTQEFWDNLEKETEGLRQEMSKDLEEV 93
DB 12 EGRGLNGFN----DNWRNAS-----LWKNIGPDAEKIKKEAPEEPI 51

QY 94 KAKVQPYLDDFOKK 107
DB 52 KEVKQENLEDLSER 65

RESULT 15
Q81R54
ID Q81R54 PRELIMINARY; PRT; 93 AA.
AC Q81R54; Q6HZC7; Q6KTB7;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=BA2204, BAS2048; ORFNames=GBAA2204;
OS Bacillus anthracis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1392;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Ames / isolate Porton;
RX MEDLINE=22608414; PubMed=12721629; DOI=10.1038/nature01586;
RA Read T.D., Peterson S.N., Tourasse N.J., Baillie L.W., Paulsen I.T.,
RA Nelson K.E., Tettelin H., Fouts D.E., Eisen J.A., Gill S.R.,
RA Holtzapple E.K., Okstad O.A., Helgason E., Rilstone J., Wu M.,
RA Kolonay J.F., Beanan M.J., Dodson R.J., Brinkac L.M., Gwinn M.L.,
RA DeBoy R.T., Madpu R., Daugherty S.C., Durkin A.S., Haft D.H.,
RA Nelson W.C., Peterson J.D., Pop M., Khouri H.M., Radune D.,
RA Benton J.L., Mahamoud Y., Jiang L., Hance I.R., Weidman J.P.,
RA Berry K.J., Plaut R.D., Wolf A.M., Watkins K.L., Nierman W.C.,
RA Hazen A., Cline R.T., Redmond C., Thwaite J.E., White O.,
RA Salzberg S.L., Thomson B., Friedlander A.M., Koehler T.M.,
RA Hanna P.C., Kolstoe A.-B., Fraser C.M.;
RT "The genome sequence of Bacillus anthracis Ames and comparison to
RT closely related bacteria.";
RL Nature 423:81-86(2003).
RN [2]
RP SEQUENCE FROM N.A.

```

```

AAQ64377
ID AAQ64377 PRELIMINARY; PRT; 76 AA.
AC AAQ64377;
DT 02-MAR-2004 (TREMBlrel. 27, Created)
DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)
DT 02-MAR-2004 (TREMBlrel. 27, Last annotation update)
DE Hypothetical protein.
GN KVP40.0307.
OS Bacteriophage KVP40.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae.
OX NCBI_TaxID=75320;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22803260; PubMed=12923095;
RA Miller E., Heidelberg J., Eise J., Nelson W., Durkin A., Ciecko A.,
RA Feldblyum T., White O., Paulsen I., Nierman W., Lee J.,
RA Szczypinski B., Fraser C.;
RT "Complete genome sequence of the broad-host-range vibriophage KVP40:
RT comparative genomics of a T4-related bacteriophage.";
RL J. Bacteriol. 185:5220-5233(2003).
RN [2]
RP SEQUENCE FROM N.A.
RA Miller E., Lee J., Szczypinski B.;
RL Submitted (Apr-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY283928; AAQ64377.1; -.
KW Hypothetical protein.
SQ SEQUENCE 76 AA; 9023 MW; 873E2003C17C60C6 CRC64;

Query Match 12.7%; Score 73; DB 2; Length 76;
Best Local Similarity 24.3%; Pred. No. 79;
Matches 18; Conservative 12; Mismatches 24; Indels 20; Gaps 2;

QY 34 EGSALGKQLNLKLLDNWDSVTSTFSKLRQLGPTQEFWDNLEKETEGLRQEMSKDLEEV 93
DB 12 EGRGLNGFN----DNWRNAS-----LWKNIGPDAEKIKKEAPEEPI 51

QY 94 KAKVQPYLDDFOKK 107
DB 52 KEVKQENLEDLSER 65

RESULT 15
Q81R54
ID Q81R54 PRELIMINARY; PRT; 93 AA.
AC Q81R54; Q6HZC7; Q6KTB7;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=BA2204, BAS2048; ORFNames=GBAA2204;
OS Bacillus anthracis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1392;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Ames / isolate Porton;
RX MEDLINE=22608414; PubMed=12721629; DOI=10.1038/nature01586;
RA Read T.D., Peterson S.N., Tourasse N.J., Baillie L.W., Paulsen I.T.,
RA Nelson K.E., Tettelin H., Fouts D.E., Eisen J.A., Gill S.R.,
RA Holtzapple E.K., Okstad O.A., Helgason E., Rilstone J., Wu M.,
RA Kolonay J.F., Beanan M.J., Dodson R.J., Brinkac L.M., Gwinn M.L.,
RA DeBoy R.T., Madpu R., Daugherty S.C., Durkin A.S., Haft D.H.,
RA Nelson W.C., Peterson J.D., Pop M., Khouri H.M., Radune D.,
RA Benton J.L., Mahamoud Y., Jiang L., Hance I.R., Weidman J.P.,
RA Berry K.J., Plaut R.D., Wolf A.M., Watkins K.L., Nierman W.C.,
RA Hazen A., Cline R.T., Redmond C., Thwaite J.E., White O.,
RA Salzberg S.L., Thomson B., Friedlander A.M., Koehler T.M.,
RA Hanna P.C., Kolstoe A.-B., Fraser C.M.;
RT "The genome sequence of Bacillus anthracis Ames and comparison to
RT closely related bacteria.";
RL Nature 423:81-86(2003).
RN [2]
RP SEQUENCE FROM N.A.

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2004, 12:24:32 ; Search time 152 Seconds
(without alignments)
257.247 Million cell updates/sec

Title: US-09-803-918A-2_COPY_25_133

Perfect score: 576
Sequence: 1 DEPPQSPWRVKDLATYYVD.....LEEVKAKVQPLDFFQKWKQ 109

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 1158598

Minimum DB seq length: 0
Maximum DB seq length: 109

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_23Sep04:.*
1: Geneseqp1980s:.*
2: Geneseqp1990s:.*
3: Geneseqp2000s:.*
4: Geneseqp2001s:.*
5: Geneseqp2002s:.*
6: Geneseqp2003as:.*
7: Geneseqp2003bs:.*
8: Geneseqp2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	197	34.2	65	6	ABP76131 Human GEN
2	162	28.1	64	5	ABP31744 Human apo
3	138	24.0	42	2	AAR20164 Apo AI po
4	137	23.8	26	7	ADE76474 ApoAI G*
5	116	20.1	21	1	AAP90956
6	112	19.4	85	8	ADP87441 Monkey ap
7	110	19.1	85	8	ADP87442 Human apo
8	108	18.8	22	1	AAP92072 Apolipop
9	108	18.8	84	2	AAY42554 Apolipop
10	108	18.8	85	8	ADP87440 Baboon ap
11	104	18.1	85	8	ADP87439 Pig apoli
12	101	17.5	85	8	ADP87438 Chicken a
13	97	16.8	19	1	AAP92073 Apolipop
14	96.5	16.8	95	6	ABP70773 Human apo
15	95	16.5	107	8	ADP87432 Rat apoli
16	95	16.5	107	8	ADP87431 Rat apoli
17	94	16.3	30	7	ADC29665 Antioxi
18	94	16.3	30	7	ADC29666 Antioxi
19	92	16.0	18	7	ADC29658 Antioxi
20	91	15.8	16	2	AAY27065 Peptide S
21	89	15.5	16	6	ABP57175 Breast ca
22	89	15.5	30	7	ADC29668 Antioxi
23	89	15.5	30	7	ADC29664 Antioxi
24	87	15.1	17	8	ADH35704 Apolipop
25	87	15.1	17	8	ADH35715 Apolipop

ALIGNMENTS

RESULT 1

ABP76131	18	7	ADC29657	Adc29657	Antioxi
ID	ABP76131	standard; protein; 65 AA.			
XX	AC				
XX	ABP76131;				
XX	21-FEB-2003	(first entry)			
XX	Human GENSET	protein SEQ ID 457.			
DE	Human GENSET	protein SEQ ID 457.			
XX					
KW	Cytostatic; antiinflammatory; nootropic; neuroprotective; cardiant;				
KW	gastrointestinal; gene therapy; GENSET; heavy metal toxicity; cancer;				
KW	inflammatory disease; immune disorder; neuromuscular; toxicity;				
KW	central nervous system; cardiovascular; gastrointestinal.				
XX					
OS	Homo sapiens.				
XX					
PN	WO200283898-A1.				
XX					
PD	24-OCT-2002.				
XX					
PF	18-APR-2001; 2001WO-IB000914.				
XX					
PR	18-APR-2001; 2001WO-IB000914.				
XX					
PA	(GEST) GENSET.				
XX					
PI	Bejanin S, Tanaka H, Dumas Milne Edwards J, Jobert S, Giordano J;				
XX	WPI; 2003-075548/07.				
XX					
DR	New GENSET polynucleotides and polypeptides, useful for treating heavy				
PT	metal toxicity, cancer, inflammatory diseases, immune disorders, and the				
PT	neuromuscular, CNS, cardiovascular or gastrointestinal effects of the				
XX	toxicity.				
PS	Claim 14; Page 535; 735pp; English.				
XX					
CC	The present invention relates to novel GENSET polynucleotides (ABZ36404-				
CC	ABZ36911) encoding polypeptides (ABP75963-ABP76368). The polynucleotides				
CC	and polypeptides are useful in screening and diagnostic assays for				
CC	abnormal GENSET expression and/or biological activity. They are also				
CC	useful for screening of compounds for treating or preventing GENSET-				
CC	related disorders, such as heavy metal toxicity, cancer, inflammatory				
CC	diseases, immune disorders, and the neuromuscular, central nervous system				
CC	(CNS), cardiovascular or gastrointestinal effects of the toxicity				
XX					
SQ	Sequence 65 AA;				

Query Match 34.2%; Score 197; DB 6; Length 65;
 Best Local Similarity 100.0%; Pred. No. 1.1e-11; Indels 0; Gaps 0;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DEPPQSPWDRVKDLATVYVDVLKDSGRDYVSQFEGSA 37
 |||||
 DB 25 DEPPQSPWDRVKDLATVYVDVLKDSGRDYVSQFEGSA 61

RESULT 2
 ABP31744
 ID ABP31744 standard; protein; 64 AA.
 XX
 AC ABP31744;
 XX
 DT 08-JUL-2002 (first entry)
 XX
 DE Human apolipoprotein-like ORF717 protein, SEQ ID NO:1434.
 XX
 KW Human; ORF; open reading frame; ORF; drug screening; diagnosis;
 KW disease monitoring; cytokine; cell proliferation; cell differentiation;
 KW immune modulation; haematopoiesis regulation; tissue growth;
 KW angiogenesis; activin; inhibin; chemotactic; chemokinetic; haemostatic;
 KW thrombolytic; tumour inhibition; bodily characteristic; fertility;
 KW behaviour; cancer; proliferative disorder; neurological disorder;
 KW cardiovascular disease; immune system disorder; organ transplantation;
 KW tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
 KW hypothyroidism; cholesterol ester storage disease; infection; vulvular;
 KW vasotrophic; antipsoriatic; antidiabetic; cytostatic; neutropic;
 KW neuroprotective; antiatherosclerotic; anticoagulant; thrombolytic;
 KW cardiant; hypotensive; antithyroid; antiinflammatory; immunomodulator;
 KW dermatological; analgesic; virucide; antibacterial; fungicide.
 XX
 OS Homo sapiens.
 XX
 PN WO200190366-A2.
 XX
 PD 29-NOV-2001.
 XX
 PF 24-MAY-2001; 2001WO-US017076.
 XX
 PR 24-MAY-2000; 2000US-0206690P.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Leach MD, Shinkets RA;
 XX
 DR WPI; 2002-106200/14.
 DR N-PSDB; ABN75770.
 XX
 PT Novel human polypeptides and polynucleotides useful for diagnosing,
 PT preventing and treating cardiovascular disease, neurodegenerative,
 PT hyperproliferative disorders and disorders related to organ
 PT transplantation.
 XX
 PS Claim 10; Page 625; 2508pp; English.
 XX

Sequences ABP31028-ABP3561 represent 4534 novel human proteins
 designated ORF (open reading frame) 1-4534, and sequences ABN75054-
 ABN79587 represent cDNAs encoding them. The invention also encompasses
 CC polypeptides at least 80% identical to the ORF1-ORF4534 (collectively
 CC referred to as ORFX) proteins, polynucleotides at least 85% identical to
 CC the ORFX nucleic acid sequences, vectors and host cells comprising ORFX
 CC polynucleotides, the recombinant production of ORFX proteins, antibodies
 CC specific for ORFX proteins, methods of detecting ORFX polynucleotides and
 CC polypeptides, methods of screening for modulators of ORFX expression or
 CC activity, and methods of screening individuals for a predisposition to an
 CC ORFX-associated disorder. The ORFX proteins of the invention have a wide
 CC range of biological activities, such as cytokine, cell proliferation,
 CC cell differentiation, immune modulation, haematopoiesis regulation,
 CC tissue growth, angiogenesis, activin or inhibin activity, chemotactic/
 CC chemokinetic activity, haemostatic activity, thrombolytic activity,

receptor/ligand, antiinflammatory activity, tumour inhibition activity,
 and antiinfective activity, and may also be involved in the determination
 of bodily characteristics, fertility and behaviour. ORFX proteins, nucleic
 acids and antibodies may be used in the treatment of cancers, other
 proliferative disorders such as psoriasis and benign tumours, neurological
 disorders such as epilepsy and Alzheimer's disease, cardiovascular diseases,
 immune system disorders, disorders related to organ transplantation,
 diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester
 storage disease, and infectious diseases caused by viral, bacterial, fungal
 and other pathogens. ORFX nucleic acids may also be used as a source of
 primers and probes, in the detection of ORFX genomic sequences or
 transcripts, in the identification and cloning of homologous sequences,
 in genetic diagnosis, and in forensic biology. The ORFX nucleic acids may
 additionally be used to produce transgenic animals which may be useful for
 studying the function and/or activity of ORFX protein, and in drug screening.
 The ORFX proteins may also be used as immunogens to generate specific
 antibodies, which are useful in the diagnosis, treatment and monitoring of
 ORFX-associated diseases

Sequence 64 AA;
 Query Match 28.1%; Score 162; DB 5; Length 64;
 Best Local Similarity 75.0%; Pred. No. 2.8e-08;
 Matches 30; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 2 EPQSPWDRVKDLATVYVDVLKDSGRDYVSQFEGSALGKQ 41
 : ||| ||||| ||||| : ||||| ||||| ||||| |||||
 DB 25 DEPPQSPWDRVKDLATVYVDVLKDSGRDYVSQFESSTLGQQ 64

RESULT 3
 AAR20164
 ID AAR20164 standard; peptide; 42 AA.
 XX
 AC AAR20164;
 XX
 DT 25-MAR-2003 (revised)
 DT 01-APR-1992 (first entry)
 XX
 DE Apo AI polypeptide (A).
 XX
 KW Monoclonal antibody; MAB AI-11; epitope; diagnosis.
 XX
 OS Synthetic.
 XX

Key Location/Qualifiers
 FT Peptide 1..28
 FT Peptide 2..28
 FT Peptide 7..28
 FT Peptide 10..28
 FT Peptide 11..42
 FT Peptide 11..28
 FT Peptide 13..28
 XX WO9118619-A.
 XX
 PD 12-DEC-1991.
 XX
 PF 07-JUN-1990; 90US-00534761.
 XX
 PR 07-JUN-1990; 90US-00534761.
 XX
 PA (SCRI) SCRIPPS CLINIC & RES.
 XX
 PI Curtiss LK, Banka CL, Bonnet DJ, Smith RS;
 DR WPI; 1992-007201/01.
 XX
 PT New Apo AI polypeptide(s) and anti-Apo antibodies - used to detect Apo AI
 PT in vascular fluid samples and increase LCAT-mediated cholesterol
 PT esterification in humans.
 XX

PS Claim 2; Page 68; 87pp; English.

XX The peptides represented in AAR20164-65 are capable of immunologically

CC mimicking an Apo AI epitope. They are useful in diagnosis and detection

CC of Apo AI in fluids and for preparing anti-Apo AI antibodies. It is

CC useful in therapeutic methods for increasing LCAT-mediated cholesterol

CC esterification in humans. The Apo AI polypeptide is selected from the

CC peptides indicated in the features and includes amino acids 13-28

CC defining a conserved native epitope on Apo AI capable of immunoreacting

CC with monoclonal antibody MAB AI-11. (Updated on 25-MAR-2003 to correct PA

CC field.)

XX

XX Sequence 42 AA;

SQ

Query Match 24.0%; Score 138; DB 2; Length 42;

Best Local Similarity 100.0%; Pred. No. 3.7e-06;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 QEMSKDLSEVKAKVQPYLDLDFQKKWQ 109

Db 1 QEMSKDLSEVKAKVQPYLDLDFQKKWQ 26

RESULT 4

ADE76474

ID ADE76474 standard; peptide; 26 AA.

XX AC ADE76474;

XX

DT 29-JAN-2004 (first entry)

XX

DE ApoAI G* amphipathic helical domain-derived peptide - SEQ ID 21.

XX

KW G-type; atherosclerosis; inflammatory; G* amphipathic helix;

KW antiarteriosclerotic; antirheumatic; antiarthritic; immunosuppressive;

KW dermatological; antiinflammatory; osteopathic; neuroprotective;

KW neurotropic; anti-HIV; virucide; antibacterial; rheumatoid arthritis;

KW systemic lupus erythematosus; polyarthritis nodosa; osteoporosis;

KW Alzheimer's disease; AIDS; infection; apoAI.

XX

OS Unidentified.

XX

XX

FT Key Location/Qualifiers

FT Misc-difference 1..26

FT /note= "Optionally D-form residue"

FT Modified-site 1

FT /note= "Optional N-terminal acetyl"

FT Modified-site 26

FT /note= "Optional C-terminal amide"

XX

XX US2003191057-A1.

XX

XX

XX 09-OCT-2003.

XX

XX 05-APR-2002; 2002US-00120508.

XX

XX 05-APR-2002; 2002US-00120508.

XX

XX (REGC) UNIV CALIFORNIA.

XX

XX Fogelman AM, Navab M;

XX

XX WPI; 2003-831662/77.

XX

XX New G-type polypeptides for preventing or ameliorating symptoms of

PT atherosclerosis or pathologies associated with an inflammatory responses

PT e.g. rheumatoid arthritis, osteoporosis, Alzheimer's disease, AIDS or

PT infections.

XX

XX Claim 7; SEQ ID NO 21; 29pp; English.

XX

XX The invention relates to a novel isolated G-type polypeptide that

CC ameliorates a symptom of atherosclerosis or other pathology associated

CC with an inflammatory response comprising an amphipathic helical peptide

CC having charged residues on the polar face of the peptide and possessing a

CC wide non-polar face. The polypeptide of the invention demonstrates

CC antiarteriosclerotic, antirheumatic, antiarthritic, immunosuppressive,

CC dermatological, antiinflammatory, osteopathic, neuroprotective,

CC neurotropic, anti-HIV, virucide and antibacterial activities. The

CC composition and methods of the invention may be useful in preventing or

CC ameliorating one or more symptoms of atherosclerosis and/or pathologies

CC characterised by an inflammatory response such as rheumatoid arthritis,

CC systemic lupus erythematosus, polyarthritis nodosa, osteoporosis,

CC Alzheimer's disease, AIDS or infection. The current sequence is that of

CC the apoAI G* amphipathic helical domain-derived peptide of the invention.

XX

XX Sequence 26 AA;

SQ

Query Match 23.8%; Score 137; DB 7; Length 26;

Best Local Similarity 100.0%; Pred. No. 2.5e-06;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 WDRVKDLATVYVDVLKDSGRDYVSQF 33

Db 1 WDRVKDLATVYVDVLKDSGRDYVSQF 26

RESULT 5

AAP90956

ID AAP90956 standard; peptide; 21 AA.

XX AC AAP90956;

XX

DT 25-MAR-2003 (revised)

DT 19-FEB-1990 (first entry)

XX

DE Apo AI epitope.

XX

XX High density lipoproteins.

KW

OS Synthetic.

XX

XX WO8909403-A.

XX

XX 05-OCT-1989.

XX

XX 27-MAR-1989; 89WO-US001262.

XX

XX 29-MAR-1988; 88US-00174698.

XX

XX (SCRI) SCRIPPS CLINIC & RES FOUND.

XX

XX Curtiss LK, Smith RS;

XX

XX WPI; 1989-309644/42.

XX

XX Apo lipoprotein peptide and monoclonal antibody against it - used to

PT assay peptide in vascular fluid to identify risk of coronary artery

PT disease.

XX

XX Claim 2; Page 41; 49pp; English.

XX

XX Apo AI contains the sequence below (Claim 1 claims AAs 1-15). There is an

CC inverse correlation between plasma levels of Apo AI and coronary artery

CC disease risk. The monoclonal antibody recognises this epitope on native

CC Apo AI. (Updated on 25-MAR-2003 to correct PA field.)

XX

XX Sequence 21 AA;

SQ

Query Match 20.1%; Score 116; DB 1; Length 21;

Best Local Similarity 100.0%; Pred. No. 0.00022;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DEPPQSPWDRVKDLATVYVDV 21

Db 1 DEPPQSPWDRVKDLATVYVDV 21

```

XX DT 09-SEP-2004 (first entry)
XX DE Human apolipoprotein A (APO-A) precursor protein.
XX KW Browseable database system; ontology; protein analysis;
XX KW gene product classification; genomic analysis; apolipoprotein A; APO-A;
XX KW human.
XX OS Homo sapiens.
XX PN WO2004053769-A2.
XX PD 24-JUN-2004.
XX PF 09-DEC-2003; 2003WO-US038935.
XX PR 09-DEC-2002; 2002US-0431879P.
XX PA (APPL-) APPLERA CORP.
XX PI Thomas PD, Kejariwal A, Campbell MJ, Mi H, Diemer K, Guo N;
PI Ladunga I, Lazareva B, Muruganujan A, Rabkin S, Vandergriff J;
PI Doremieux O;
XX PI Doremieux O;
XX DR WPI; 2004-480967/45.
XX XX
XX PT Browseable database system, for e.g. analyzing protein sequences, and
XX PT predicting a biological role, comprises a datastore, an ontology of
XX PT categories of biological functions, an input receptive, a recognizer, and
XX PT an output.
XX PS Disclosure; SEQ ID NO 49; 113pp; English.
XX CC The invention relates to a browsable database system for use with
XX CC biological information, comprises a datastore, an ontology of categories
XX CC of biological functions, an input receptive, a recognizer and an output.
XX CC The invention also provides a method of operation for use with a
XX CC browsable biological database system. The browsable database system and
XX CC methods are useful in analysing protein sequences, classifying gene
XX CC products, predicting biological role for pathway building, enhancing
XX CC interpretation of expression information, providing protein function and
XX CC in facilitating comparative genomic analysis. The present sequence is
XX CC human apolipoprotein A (APO-A) IV precursor protein. This sequence is
XX CC used to illustrate the method of the invention.
XX SQ Sequence 85 AA;
XX Query Match 19.4%; Score 112; DB 8; Length 85;
XX Best Local Similarity 33.3%; Pred. No. 0.0031;
XX Matches 24; Conservative 19; Mismatches 29; Indels 0; Gaps 0;

QY 36 SALGKQLNLKLDNWDVSTFTSKLREQLGVPVTQEFWDNLEKETEGLRQEMSKDLEEVKA 95
Db 1 SELTQQLNALFQDKLGEVNTYAGDLQKLVPPATELHERLAKDSEKLEKEIRKELEEVRA 60

QY 96 KVQPYLDDFQKK 107
Db 61 RLLPHANEVSQK 72

RESULT 7
ADP87442
ID ADP87442 standard; protein; 85 AA.
XX AC ADP87442;
XX DT 09-SEP-2004 (first entry)
XX DE Monkey apolipoprotein A (APO-A) precursor protein.
XX KW Browseable database system; ontology; protein analysis;
XX KW gene product classification; genomic analysis; apolipoprotein A; APO-A;
XX KW monkey.
XX OS Macaca fascicularis.
XX PN WO2004053769-A2.
XX PD 24-JUN-2004.
XX PF 09-DEC-2003; 2003WO-US038935.
XX PR 09-DEC-2002; 2002US-0431879P.
XX PA (APPL-) APPLERA CORP.
XX PI Thomas PD, Kejariwal A, Campbell MJ, Mi H, Diemer K, Guo N;
PI Ladunga I, Lazareva B, Muruganujan A, Rabkin S, Vandergriff J;
PI Doremieux O;
XX PI Doremieux O;
XX DR WPI; 2004-480967/45.
XX XX
XX PT Browseable database system, for e.g. analyzing protein sequences, and
XX PT predicting a biological role, comprises a datastore, an ontology of
XX PT categories of biological functions, an input receptive, a recognizer, and
XX PT an output.
XX PS Disclosure; SEQ ID NO 48; 113pp; English.
XX CC The invention relates to a browsable database system for use with
XX CC biological information, comprises a datastore, an ontology of categories
XX CC of biological functions, an input receptive, a recognizer and an output.
XX CC The invention also provides a method of operation for use with a
XX CC browsable biological database system. The browsable database system and
XX CC methods are useful in analysing protein sequences, classifying gene
XX CC products, predicting biological role for pathway building, enhancing
XX CC interpretation of expression information, providing protein function and
XX CC in facilitating comparative genomic analysis. The present sequence is
XX CC monkey apolipoprotein A (APO-A) IV precursor protein. This sequence is
XX CC used to illustrate the method of the invention.
XX SQ Sequence 85 AA;
XX Query Match 19.4%; Score 112; DB 8; Length 85;
XX Best Local Similarity 33.3%; Pred. No. 0.0031;
XX Matches 24; Conservative 19; Mismatches 29; Indels 0; Gaps 0;

QY 36 SALGKQLNLKLDNWDVSTFTSKLREQLGVPVTQEFWDNLEKETEGLRQEMSKDLEEVKA 95
Db 1 SELTQQLNALFQDKLGEVNTYAGDLQKLVPPATELHERLAKDSEKLEKEIRKELEEVRA 60

QY 96 KVQPYLDDFQKK 107
Db 61 RLLPHANEVSQK 72

RESULT 8
AAP92072
ID AAP92072 standard; peptide; 22 AA.
XX AC AAP92072;
XX DT 25-MAR-2003 (revised)
XX DT 31-OCT-2002 (revised)
XX DT 10-APR-1990 (first entry)
XX DE Apolipoprotein AI (Apo AI) polypeptide.
XX XX
```

```

XX DT 09-SEP-2004 (first entry)
XX DE Human apolipoprotein A (APO-A) precursor protein.
XX KW Browseable database system; ontology; protein analysis;
XX KW gene product classification; genomic analysis; apolipoprotein A; APO-A;
XX KW human.
XX OS Homo sapiens.
XX PN WO2004053769-A2.
XX PD 24-JUN-2004.
XX PF 09-DEC-2003; 2003WO-US038935.
XX PR 09-DEC-2002; 2002US-0431879P.
XX PA (APPL-) APPLERA CORP.
XX PI Thomas PD, Kejariwal A, Campbell MJ, Mi H, Diemer K, Guo N;
PI Ladunga I, Lazareva B, Muruganujan A, Rabkin S, Vandergriff J;
PI Doremieux O;
XX PI Doremieux O;
XX DR WPI; 2004-480967/45.
XX XX
XX PT Browseable database system, for e.g. analyzing protein sequences, and
XX PT predicting a biological role, comprises a datastore, an ontology of
XX PT categories of biological functions, an input receptive, a recognizer, and
XX PT an output.
XX PS Disclosure; SEQ ID NO 49; 113pp; English.
XX CC The invention relates to a browsable database system for use with
XX CC biological information, comprises a datastore, an ontology of categories
XX CC of biological functions, an input receptive, a recognizer and an output.
XX CC The invention also provides a method of operation for use with a
XX CC browsable biological database system. The browsable database system and
XX CC methods are useful in analysing protein sequences, classifying gene
XX CC products, predicting biological role for pathway building, enhancing
XX CC interpretation of expression information, providing protein function and
XX CC in facilitating comparative genomic analysis. The present sequence is
XX CC human apolipoprotein A (APO-A) IV precursor protein. This sequence is
XX CC used to illustrate the method of the invention.
XX SQ Sequence 85 AA;
XX Query Match 19.1%; Score 110; DB 8; Length 85;
XX Best Local Similarity 31.9%; Pred. No. 0.0049;
XX Matches 23; Conservative 20; Mismatches 29; Indels 0; Gaps 0;

QY 36 SALGKQLNLKLDNWDVSTFTSKLREQLGVPVTQEFWDNLEKETEGLRQEMSKDLEEVKA 95
Db 1 SELTQQLNALFQDKLGEVNTYAGDLQKLVPPATELHERLAKDSEKLEKEIRKELEEVRA 60

QY 96 KVQPYLDDFQKK 107
Db 61 RLLPHANEVSQK 72

RESULT 8
AAP92072
ID AAP92072 standard; peptide; 22 AA.
XX AC AAP92072;
XX DT 25-MAR-2003 (revised)
XX DT 31-OCT-2002 (revised)
XX DT 10-APR-1990 (first entry)
XX DE Apolipoprotein AI (Apo AI) polypeptide.
XX XX
```


KW Apolipoprotein AI; Apo AI; high-density lipoprotein; HDL;
coronary artery disease; CAD; Apo AI epitope.

XX Homo sapiens.

PH Key Location/Qualifiers
FT Peptide 1. .22 /label= AI90-111
FT Peptide 1. .22 /label= AI90-111
FT Peptide 1.16 /label= AI90-105
FT Peptide 4. .12 /label= AI93-101
FT Peptide 6. .16 /label= AI95-105
FT Peptide 7.12 /label= AI96-101
FT Peptide 11. .16 /label= AI100-105
FT Peptide 12. .22 /label= AI101-111

XX WO8904486-A.

XX 18-MAY-1989.

XX 02-NOV-1988; 88WO-US003903.

XX 03-NOV-1987; 87US-00116248.

XX (SCRI) SCRIPPS CLINIC & RES FOUND.

XX Curtiss LK, Smith RR;

XX WPI; 1989-165740/22.

XX New monoclonal antibody and polypeptide antigens - directed against APO AI-HDL epitope, useful in diagnostic assays.

XX Claim 2; Table 1-2; 62pp; English.

XX AI90-111 and smaller peptides contained within this sequence. AI95-105, CC AI190-105 and AI190-111 are specifically claimed. Such Apo AI peptides are capable of immunologically mimicking a native conserved Apo AI epitope. They may be helpful in the diagnosis of risk of coronary heart disease. (Updated on 31-OCT-2002 to add missing OS field.) (Updated on 25 -MAR-2003 to correct PR field.) (Updated on 25-MAR-2003 to correct PR field.)

XX Sequence 22 AA;

Query Match 18.8%; Score 108; DB 1; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 90 LEEVKAKVQPYLDDFQKKWQ 109
Db 1 LEEVKAKVQPYLDDFQKKWQ 20

RESULT 9

AAY42554

ID AAY42554 standard; peptide; 84 AA.

XX AAY42554;

XX 20-DEC-1999 (first entry)

XX Apolipoprotein A-IV derived lipid oxidation suppressant peptide #2.

XX Cholesterol; cardiovascular disease; heart disease; atherosclerosis;
KW lipoprotein; angina; myocardial infarction; stroke; thrombosis;

KW antioxidant; hypolipidaemic; apolipoprotein.

XX Synthetic.

XX Homo sapiens.

PH Key Location/Qualifiers

FT Modified-site 84 /note= "Optionally has a C-terminal amide"

XX WO9950286-A2.

XX 07-OCT-1999.

XX 25-MAR-1999; 99WO-US006580.

XX 31-MAR-1998; 98US-0080131P.

XX (UYCI-) UNIV CINCINNATI.

XX Hui DY, Tso P;

XX WPI; 1999-580739/49.

XX Treating conditions associated with lipid oxidation or preventing oxidation in lipid-containing food, lipid containing pharmaceuticals or cosmetic or dermatological compositions.

XX Claim 4; Page 53; 73pp; English.

XX This sequence represents a peptide (#2) derived from apolipoprotein (apo) A-IV with lipid oxidation inhibitory activity. Lipid oxidation plays a role in the development of atherosclerosis, a main cause of coronary heart disease. Atherosclerosis is thought to begin with local injury to the arterial endothelium, followed by proliferation of arterial smooth muscle cells, along with deposition of lipid and accumulation of foam cells in the lesion. As the atherosclerotic plaque develops, it progressively occludes more and more blood vessel and can eventually lead to ischemia or infarction. Both the plasma concentration and qualitative characteristics of low density lipoproteins (LDL) are risk factors in atherogenesis. Oxidation causes important changes in the primary structure of the main LDL apolipoprotein, apolipoprotein B100 (apo B-100). These changes, by helping LDL absorption by macrophages, causes the intracellular accumulation of esters of cholesterol and the formation of foam cells, with subsequent development of the atherosclerotic plaque. The apo A-IV derived peptides can be used for inhibiting lipid oxidation. The peptides can be used for treating conditions associated with lipid oxidation. For example, they can be used for treating or inhibiting the progression of atherosclerosis. They can also be used for preventing oxidation in lipid-containing foods, lipid-containing pharmaceuticals or cosmetic or dermatological compositions. As these peptides comprise specific portions of the native apo A-IV protein, there should be no immunogenicity problems with their administration to humans

XX Sequence 84 AA;

Query Match 18.8%; Score 108; DB 2; Length 84;
Best Local Similarity 27.8%; Pred. No. 0.0076;
Matches 25; Conservative 22; Mismatches 35; Indels 8; Gaps 1;

OY 8 WDRVKDLATVYVDVLKDSGRDYVSQFEGSALGKQLNLKLLDNWDSVTFSKRLQGLGPV 67

Db 2 WD-----YFTQLSNNAKEAVEQLQKTDVTQQLNTLFDQKLGNINTYADDLQNKLVFP 53

OY 68 TQEFWDNLEKETEGLRQEMSKDLEEVKAKV 97

Db 54 AVQLSGHLTKETERVREEIQKELDLRAMV 83

RESULT 10

ADP87440

ID ADP87440 standard; protein; 85 AA.

XX ADP87440;

```
XX 09-SEP-2004 (first entry)
DT Baboon apolipoprotein A (APO-A) IV precursor protein.
XX
XX
DE Browseable database system; ontology; protein analysis;
XX gene product classification; genomic analysis; apolipoprotein A; APO-A;
KW Baboon.
XX
XX
OS Papio anubis.
XX
XX WO2004053769-A2.
PN
XX
XX 24-JUN-2004.
PD
XX
XX 09-DEC-2003; 2003WO-US038935.
PF
XX
XX 09-DEC-2002; 2002US-0431879P.
PR
XX
XX (APPL-) APPLERA CORP.
PA
XX Thomas PD, Kejariwal A, Campbell MJ, Mi H, Diemer K, Guo N;
PI Ladunga I, Lazareva B, Muruganujan A, Rabkin S, Vandergriff J;
PI Doremieux O;
XX
XX WPI; 2004-480967/45.
DR
XX
XX Browseable database system, for e.g. analyzing protein sequences, and
XX predicting a biological role, comprises a database, an ontology of
XX categories of biological functions, an input receptive, a recognizer, and
XX an output.
XX
XX Disclosure; SEQ ID NO 47; 113pp; English.
XX
XX The invention relates to a browsable database system for use with
XX biological information, comprises a database, an ontology of categories
XX of biological functions, an input receptive, a recognizer and an output.
XX The invention also provides a method of operation for use with a
XX browsable biological database system. The browsable database system and
XX methods are useful in analysing protein sequences, classifying gene
XX products, predicting biological role for pathway building, enhancing
XX interpretation of expression information, providing protein function and
XX in facilitating comparative genomic analysis. The present sequence is
XX baboon apolipoprotein A (APO-A) IV precursor protein. This sequence is
XX used to illustrate the method of the invention.
XX
XX Sequence 85 AA;
QY Query Match 18.8%; Score 108; DB 8; Length 85;
Best Local Similarity 31.9%; Pred. No. 0.0077;
Matches 23; Conservative 20; Mismatches 29; Indels 0; Gaps 0;
QY 36 SALGKQLNLKLLDNWDSVTSTFSKLRQGLGPTVQTFWNLKETEGLRQEMSKDLSEVKA 95
Db 1 SELTQQLNALFQDKLGEVNTYAGDLQKLVPPFATLHERLAKDSKKLKEIRKELEEVRA 60
QY 96 KVQPYLDDFQKK 107
Db 61 RLLPHANEVSQK 72
RESULT 11
ADP87439
ID ADP87439 standard; protein; 85 AA.
XX
XX ADP87439;
XX
XX 09-SEP-2004 (first entry)
XX
XX Pig apolipoprotein A (APO-A) IV precursor protein.
DE
XX Browseable database system; ontology; protein analysis;
KW gene product classification; genomic analysis; apolipoprotein A; APO-A;
KW
```

```
KW pig.
XX
XX Sus scrofa.
XX
XX WO2004053769-A2.
PN
XX
XX 24-JUN-2004.
PD
XX
XX 09-DEC-2003; 2003WO-US038935.
PF
XX
XX 09-DEC-2002; 2002US-0431879P.
PR
XX
XX (APPL-) APPLERA CORP.
PA
XX Thomas PD, Kejariwal A, Campbell MJ, Mi H, Diemer K, Guo N;
PI Ladunga I, Lazareva B, Muruganujan A, Rabkin S, Vandergriff J;
PI Doremieux O;
XX
XX WPI; 2004-480967/45.
DR
XX
XX Browseable database system, for e.g. analyzing protein sequences, and
XX predicting a biological role, comprises a database, an ontology of
XX categories of biological functions, an input receptive, a recognizer, and
XX an output.
XX
XX Disclosure; SEQ ID NO 46; 113pp; English.
XX
XX The invention relates to a browsable database system for use with
XX biological information, comprises a database, an ontology of categories
XX of biological functions, an input receptive, a recognizer and an output.
XX The invention also provides a method of operation for use with a
XX browsable biological database system. The browsable database system and
XX methods are useful in analysing protein sequences, classifying gene
XX products, predicting biological role for pathway building, enhancing
XX interpretation of expression information, providing protein function and
XX in facilitating comparative genomic analysis. The present sequence is pig
XX apolipoprotein A (APO-A) IV precursor protein. This sequence is used to
XX illustrate the method of the invention.
XX
XX Sequence 85 AA;
QY Query Match 18.1%; Score 104; DB 8; Length 85;
Best Local Similarity 30.6%; Pred. No. 0.019;
Matches 22; Conservative 20; Mismatches 30; Indels 0; Gaps 0;
QY 36 SALGKQLNLKLLDNWDSVTSTFSKLRQGLGPTVQTFWNLKETEGLRQEMSKDLSEVKA 95
Db 1 SELTQQLNLTLFQDKLGEVNTYTDLQKLVPPFATLHERLTKDSEKLEIRRELEELA 60
QY 96 KVQPYLDDFQKK 107
Db 61 RLLPHATEVSQK 72
RESULT 12
ADP87438
ID ADP87438 standard; protein; 85 AA.
XX
XX ADP87438;
XX
XX 09-SEP-2004 (first entry)
XX
XX Chicken apolipoprotein A (APO-A) IV protein.
DE
XX Browseable database system; ontology; protein analysis;
KW gene product classification; genomic analysis; apolipoprotein A; APO-A;
KW chicken.
XX
XX Gallus gallus.
XX
XX WO2004053769-A2.
PN
XX
XX 24-JUN-2004.
```


PS Claim 1; Page 6; 45pp; French.

XX The present sequence is a peptide fragment of human Apolipoprotein AIV

CC Related Protein (AA4RP). The peptide fragment is useful for raising

CC specific antibodies (Ab) against AA4RP, so are useful in immunoassays for

CC detecting or determining AA4RP, especially to identify subjects at risk

CC of developing disorders of lipid metabolism (particularly cardiovascular

CC diseases associated with hyperlipidemia, e.g. coronary disease). Ab are

CC also useful for protein purification and, optionally when conjugated to a

CC toxin or pharmaceutical, for treating the specified diseases

XX Sequence 95 AA;

SQ

Query Match 16.8%; Score 96.5; DB 6; Length 95;

Best Local Similarity 29.5%; Pred. No. 0.12; Mismatches 21; Indels 3; Gaps 2;

Matches 23; Conservative 21;

QY 24 DSGRDYVSQFEGSALGKQLNLKLDNWDVSTFSLREQLGPGVTQEFWDNLEKETEGLR 83

Db 18 DKGR--VEQIHQQKWARE-PATLKDSLEQDLNNMKFLEKLRLSLGSEAPRLPDQPVGMR 74

QY 84 QEMSKDLEEVKAKVPYL 101

Db 75 RQLQEELEEVKARLPQM 92

RESULT 15

ADP87432

ID ADP87432 standard; protein; 107 AA.

XX

AC ADP87432;

XX

DT 09-SEP-2004 (first entry)

XX

DE Rat apolipoprotein C (APO-C) IV protein.

XX

XX Browseable database system; ontology; protein analysis;

KW gene product classification; genomic analysis; apolipoprotein C; APO-C;

KW rat.

XX

OS Rattus norvegicus.

XX

XX WO2004053769-A2.

PN

XX

PD 24-JUN-2004.

XX

XX 09-DEC-2003; 2003WO-US038935.

PF

XX

XX 09-DEC-2002; 2002US-0431879P.

PR

XX

XX (APPL-) APPLERA CORP.

PA

XX

PI Thomas PD, Kejariwal A, Campbell MJ, Mi H, Diemer K, Guo N;

PI Ladunga I, Lazareva B, Muruganujan A, Rabkin S, Vandergriff J;

PI Doremieux O;

XX

XX WPI; 2004-480967/45.

DR

XX

XX Browseable database system, for e.g. analyzing protein sequences, and

PT predicting a biological role, comprises a datastore, an ontology of

PT categories of biological functions, an input receptive, a recognizer, and

PT an output.

XX

PS Disclosure; SEQ ID NO 39; 113pp; English.

XX

XX The invention relates to a browsable database system for use with

CC biological information, comprises a datastore, an ontology of categories

CC of biological functions, an input receptive, a recognizer and an output.

CC The invention also provides a method of operation for use with a

CC browsable biological database system. The browsable database system and

CC methods are useful in analysing protein sequences, classifying gene

CC products, predicting biological role for pathway building, enhancing

CC interpretation of expression information, providing protein function and

CC in facilitating comparative genomic analysis. The present sequence is rat

CC apolipoprotein C (APO-C) IV protein. This sequence is used to illustrate

CC the method of the invention.

XX

SQ Sequence 107 AA;

Query Match 16.5%; Score 95; DB 8; Length 107;

Best Local Similarity 29.3%; Pred. No. 0.19; Mismatches 27; Indels 8; Gaps 1;

Matches 22; Conservative 18;

QY 40 KQLNLKLLDNWDVSTFSLREQLGPGVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVOP 99

Db 5 QQLNTLFDQDKLGNINTYADDLQNKLVPPFVQLSGHLTKETERVREIQKELEDLRANMMP 64

QY 100 YL-----DDFQK 106

Db 65 HANKVSQMFQGNVQK 79

Search completed: December 21, 2004, 12:37:24

Job time : 156 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2004, 12:14:35 ; Search time 144 Seconds
(without alignments)
270.861 Million cell updates/sec

Title: US-09-803-918A-2_COPY_25_133

Perfect score: 576

Sequence: 1 DEBPQSPWDRVKDLATVYVD.....LEEVKAKVQPYLDDFOKKWQ 109

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1589859 seqs, 357834939 residues

Total number of hits satisfying chosen parameters: 739951

Minimum DB seq length: 0

Maximum DB seq length: 109

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	332.5	57.7	93	15 US-10-038-854-403	Sequence 403, App
2	332.5	57.7	93	15 US-10-038-854-405	Sequence 405, App
3	318	55.2	86	15 US-10-424-599-276814	Sequence 276814,
4	227	39.4	43	15 US-10-465-789A-99	Sequence 89, Appl
5	162	28.1	64	11 US-09-864-408A-1434	Sequence 1434, Ap
6	137	23.8	26	14 US-10-120-508-21	Sequence 21, Appl
7	118	20.5	22	15 US-10-465-789A-50	Sequence 50, Appl
8	112	19.4	22	15 US-10-465-789A-49	Sequence 49, Appl
9	96.5	16.8	95	17 US-10-487-096-2	Sequence 2, Appli
10	94	16.3	30	14 US-10-142-238A-40	Sequence 40, Appl
11	94	16.3	30	14 US-10-142-238A-41	Sequence 41, Appl
12	92	16.0	18	14 US-10-142-238A-33	Sequence 33, Appl
13	89	15.5	16	17 US-10-700-340-37	Sequence 37, Appl

14	89	15.5	30	14	US-10-142-238A-39	Sequence 39, Appl
15	89	15.5	30	14	US-10-142-238A-43	Sequence 43, Appl
16	87	15.1	17	15	US-10-601-100-44	Sequence 44, Appl
17	87	15.1	18	14	US-10-142-238A-32	Sequence 32, Appl
18	87	15.1	18	14	US-10-142-238A-35	Sequence 35, Appl
19	86	14.9	30	14	US-10-142-238A-42	Sequence 42, Appl
20	86	14.9	30	14	US-10-142-238A-46	Sequence 46, Appl
21	85	14.8	30	14	US-10-142-238A-44	Sequence 44, Appl
22	85	14.8	30	14	US-10-142-238A-45	Sequence 45, Appl
23	84	14.6	18	14	US-10-142-238A-34	Sequence 34, Appl
24	84	14.6	18	14	US-10-142-238A-38	Sequence 38, Appl
25	83	14.4	18	14	US-10-142-238A-36	Sequence 36, Appl
26	83	14.4	18	14	US-10-142-238A-37	Sequence 37, Appl
27	74	12.8	14	15	US-10-601-100-43	Sequence 43, Appl
28	67.5	11.7	72	16	US-10-767-701-37365	Sequence 37365, A
29	67	11.6	21	14	US-10-076-047A-34	Sequence 34, Appl
30	67	11.6	22	15	US-10-465-789A-52	Sequence 52, Appl
31	66	11.5	13	14	US-10-014-340-123	Sequence 123, App
32	66	11.5	13	15	US-10-264-309-76	Sequence 76, Appl
33	66	11.5	13	15	US-10-601-100-36	Sequence 36, Appl
34	65.5	11.4	99	15	US-10-674-755-16	Sequence 16, Appl
35	64	11.1	62	9	US-09-864-761-40424	Sequence 40424, A
36	63	10.9	10	14	US-10-033-741-11	Sequence 11, Appl
37	63	10.9	10	14	US-10-014-340-79	Sequence 79, Appl
38	63	10.9	10	15	US-10-264-309-52	Sequence 52, Appl
39	63	10.9	10	15	US-10-601-100-55	Sequence 55, Appl
40	63	10.9	10	17	US-10-700-340-21	Sequence 21, Appl
41	62.5	10.9	97	17	US-10-425-115-300637	Sequence 300637,
42	62	10.8	80	15	US-10-424-599-197985	Sequence 197985,
43	62	10.8	108	15	US-10-424-599-243777	Sequence 243777,
44	61.5	10.7	99	15	US-10-674-755-11	Sequence 11, Appl
45	61	10.6	100	15	US-10-674-755-12	Sequence 12, Appl

ALIGNMENTS

RESULT 1

US-10-038-854-403
; Sequence 403, Application US/10038854
; Publication No. US20040022781A1
; GENERAL INFORMATION:
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Li, Li
; APPLICANT: Wolenc, Adam R
; APPLICANT: Vernet, Corine
; APPLICANT: Eisen, Andrew J
; APPLICANT: Liu, Xiaohong
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Shinkets, Richard A
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spaderna, Steven K
; APPLICANT: Gorman, Linda
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Gangolli, Bsha A
; APPLICANT: Guo, Xiaojia S
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Rastelli, Luca
; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc
; APPLICANT: Burgess, Catherine E
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gunther, Erik
; APPLICANT: Smithson, Glennnda
; APPLICANT: Millet, Isabelle
; APPLICANT: MacDougall, John R
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-230
; CURRENT APPLICATION NUMBER: US/10/038,854
; CURRENT FILING DATE: 2003-01-22

APPLICANT: MacDougall, John R
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-230
CURRENT APPLICATION NUMBER: US/10/038,854
CURRENT FILING DATE: 2003-01-22
PRIOR APPLICATION NUMBER: 60/258,928
PRIOR FILING DATE: 2000-12-29
PRIOR APPLICATION NUMBER: 60/259,415
PRIOR FILING DATE: 2001-01-02
PRIOR APPLICATION NUMBER: 60/259,785
PRIOR FILING DATE: 2001-01-04
PRIOR APPLICATION NUMBER: 60/269,814
PRIOR FILING DATE: 2001-02-20
PRIOR APPLICATION NUMBER: 60/279,832
PRIOR FILING DATE: 2001-03-29
PRIOR APPLICATION NUMBER: 60/279,833
PRIOR FILING DATE: 2001-03-29
PRIOR APPLICATION NUMBER: 60/279,863
PRIOR FILING DATE: 2001-03-29
PRIOR APPLICATION NUMBER: 60/283,889
PRIOR FILING DATE: 2001-04-13
PRIOR APPLICATION NUMBER: 60/284,447
PRIOR FILING DATE: 2001-04-18
PRIOR APPLICATION NUMBER: 60/286,683
PRIOR FILING DATE: 2001-04-25
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 411
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 403
LENGTH: 93
TYPE: PRT
ORGANISM: Homo sapiens
US-10-038-854-403

Query Match 57.7%; Score 332.5; DB 15; Length 93;
Best Local Similarity 71.9%; Pred. No. 4e-25;
Matches 69; Conservative 0; Mismatches 0; Indels 27; Gaps 1;

QY 14 LATVYVDVLKDSGRDYSQFEGSALGKQLNLKLLDNWDSVTSTFSKLRQLGPGVTQEFWD 73
Db 1 LATVYVDVLK-----DSVTSTFSKLRQLGPGVTQEFWD 33
QY 74 NLEKETEGLRQEMSKDLEEVKAKVQPYLDDFOKKWQ 109
Db 34 NLEKETEGLRQEMSKDLEEVKAKVQPYLDDFOKKWQ 69

RESULT 2
US-10-038-854-405
Sequence 405, Application US/10038854
Publication No. US20040022781A1
GENERAL INFORMATION:
APPLICANT: Spytek, Kimberly A
APPLICANT: Li, Li
APPLICANT: Wolenc, Adam R
APPLICANT: Vernet, Corine
APPLICANT: Eissen, Andrew J
APPLICANT: Liu, Xiaohong
APPLICANT: Malyankar, Uriel M
APPLICANT: Shimkets, Richard A
APPLICANT: Tchernev, Velizar
APPLICANT: Spaderna, Steven K
APPLICANT: Gorman, Linda
APPLICANT: Kekuda, Ramesh
APPLICANT: Patturajan, Meera
APPLICANT: Gusev, Vladimir Y
APPLICANT: Gangolli, Baha A
APPLICANT: Guo, Xiaojia S
APPLICANT: Shenoy, Suresh G
APPLICANT: Rastelli, Luca
APPLICANT: Casman, Stacie J
APPLICANT: Boldog, Ferenc
APPLICANT: Burgess, Catherine E
APPLICANT: Edinger, Shlomit R
APPLICANT: Ellerman, Karen
APPLICANT: Gunther, Erik
APPLICANT: Smithson, Glenda
APPLICANT: Millet, Isabelle

APPLICANT: MacDougall, John R
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-230
CURRENT APPLICATION NUMBER: US/10/038,854
CURRENT FILING DATE: 2003-01-22
PRIOR APPLICATION NUMBER: 60/258,928
PRIOR FILING DATE: 2000-12-29
PRIOR APPLICATION NUMBER: 60/259,415
PRIOR FILING DATE: 2001-01-02
PRIOR APPLICATION NUMBER: 60/259,785
PRIOR FILING DATE: 2001-01-04
PRIOR APPLICATION NUMBER: 60/269,814
PRIOR FILING DATE: 2001-02-20
PRIOR APPLICATION NUMBER: 60/279,832
PRIOR FILING DATE: 2001-03-29
PRIOR APPLICATION NUMBER: 60/279,833
PRIOR FILING DATE: 2001-03-29
PRIOR APPLICATION NUMBER: 60/279,863
PRIOR FILING DATE: 2001-03-29
PRIOR APPLICATION NUMBER: 60/283,889
PRIOR FILING DATE: 2001-04-13
PRIOR APPLICATION NUMBER: 60/284,447
PRIOR FILING DATE: 2001-04-18
PRIOR APPLICATION NUMBER: 60/286,683
PRIOR FILING DATE: 2001-04-25
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 411
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 405
LENGTH: 93
TYPE: PRT
ORGANISM: Homo sapiens
US-10-038-854-405

Query Match 57.7%; Score 332.5; DB 15; Length 93;
Best Local Similarity 71.9%; Pred. No. 4e-25;
Matches 69; Conservative 0; Mismatches 0; Indels 27; Gaps 1;
QY 14 LATVYVDVLKDSGRDYSQFEGSALGKQLNLKLLDNWDSVTSTFSKLRQLGPGVTQEFWD 73
Db 1 LATVYVDVLK-----DSVTSTFSKLRQLGPGVTQEFWD 33
QY 74 NLEKETEGLRQEMSKDLEEVKAKVQPYLDDFOKKWQ 109
Db 34 NLEKETEGLRQEMSKDLEEVKAKVQPYLDDFOKKWQ 69

RESULT 3
US-10-424-599-276814
Sequence 276814, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 276814
LENGTH: 86
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_91983C.1.pap
US-10-424-599-276814

Query Match 55.2%; Score 318; DB 15; Length 86;
Best Local Similarity 98.4%; Pred. No. 9.8e-24;
Matches 61; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	2	EP	QSP	WDR	VK	DL	AT	VY	VD	VL	KD	SG	RD	YV	SQ	FE	GS	AL	GK	Q	41
	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
Db	25	DE	POS	QW	DR	VK	OF	AT <td>VY</td> <td>VD <td>AV</td> <td>KD</td> <td>SG</td> <td>RN</td> <td>YV</td> <td>SQ <td>FE</td> <td>SS</td> <td>TL</td> <td>GO <td>64</td> </td></td></td>	VY	VD <td>AV</td> <td>KD</td> <td>SG</td> <td>RN</td> <td>YV</td> <td>SQ <td>FE</td> <td>SS</td> <td>TL</td> <td>GO <td>64</td> </td></td>	AV	KD	SG	RN	YV	SQ <td>FE</td> <td>SS</td> <td>TL</td> <td>GO <td>64</td> </td>	FE	SS	TL	GO <td>64</td>	64

RESULT 6
US-10-120-508-21
; Sequence 21, Application US/10120508
; Publication No. US20030191057A1

```

1  APPLICANT: FOGELMAN, ALAN
2  APPLICANT: NAVAB, MOHAWAD
3  TITLE OF INVENTION: G-TYPE PEPTIDES TO AMELIORATE ATHEROSCLEROSIS
4  FILE REFERENCE: 407T-301100US
5  CURRENT APPLICATION NUMBER: US/10/120,508
6  CURRENT FILING DATE: 2002-04-05
7  NUMBER OF SEQ ID NOS: 32
8  SOFTWARE: PatentIn version 3.0
9  SEQ ID NO 21
10 LENGTH: 26
11 TYPE: PRT
12 ORGANISM: Artificial Sequence
13 FEATURE:
14 OTHER INFORMATION: Synthetic D peptide.
15 US-10-120-508-21

```

```
Query Match      23.8%; Score 137; DB 14; Length 26;
Best Local Similarity 100.0%; Pred. No. 1.6e-06;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

8 WDRVKDLATVYVDVLKDSGRDYVSQF 33
|||||
1 WDRVKDLATVYVDVLKDSGRDYVSQF 26

RESULT 7
US-10-465-789A-50
; Sequence 50, Application US/10465789A
; Publication No. US20040053384A1
; GENERAL INFORMATION:

```
Query Match          39.4%; Score 227; DB 15; Length 43;
Best Local Similarity 100.0%; Pred. No. 3.9e-15;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

1 DEPPQSPWDRVKDLATVYVDVLKDSGRDYVSQPEGSALGKQLN 43

RESULT 5
 S-09-864-408A-1434
 Sequence 1434, Application US/09864408A
 Publication No. US20040009474A1
 GENERAL INFORMATION:
 APPLICANT: Leach, Martin D.
 APPLICANT: Shimkets, Richard A.
 TITLE OF INVENTION: No. US20040009474A1el Human Polynucleotides and Polypeptides Encoded by the Same
 FILE REFERENCE: 21402-012
 CURRENT APPLICATION NUMBER: US/09/864, 408A
 CURRENT FILING DATE: 2001-05-24
 PRIOR APPLICATION NUMBER: 60/206,690
 PRIOR FILING DATE: 2000-05-24
 NUMBER OF SEQ ID NOS: 9068
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 1434
 LENGTH: 64
 TYPE: PRT
 ORGANISM: Homo sapiens
 S-09-864-408A-1434

Query Match 28.1%; Score 162; DB 11; Length 64;
Best Local Similarity 75.0%; Pred. No. 1.7e-08;
Matches 30: Conservative 4; Mismatches 6; Indels 0; Gaps 0;

```
Query Match      20.5%; Score 118; DB 15; Length 22;
Best Local Similarity 100.0%; Pred. No. 9.8e-05;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

66 PVTQEFWDNLEKETEGLRQEMS 87
1 PVTQEFWDNLEKETEGLRQEMS 22

```
RESULT 8
US-10-465-789A-49
; Sequence 49, Application US/10465789A
; Publication No. US20040053384A1
; GENERAL INFORMATION:
; APPLICANT: Sligar, Stephen G
; APPLICANT: Bayburt, Timothy H
; APPLICANT: Schuler, Mary A
; APPLICANT: Civjan, Natanya R
; APPLICANT: Yelena V. Grinkova
; APPLICANT: Ilia G. Denisov
; TITLE OF INVENTION: Membrane Scaffold Proteins
; FILE REFERENCE: 87-00A
; CURRENT APPLICATION NUMBER: US/10/465,789A
; CURRENT FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: 09/990,087
; PRIOR FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 60/252,233
; PRIOR FILING DATE: 2000-11-20
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 49
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: artificial sequence of Helix 1
US-10-465-789A-49

Query Match          19.4%; Score 112; DB 15; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.00038;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 LKLLDNWDSVTSTFSKLRQLG 65
| | | | | | | | | | | | | | | | | | | |
Db 1 LKLLDNWDSVTSTFSKLRQLG 22

RESULT 9
US-10-487-096-2
; Sequence 2, Application US/10487096
; Publication No. US20040197823A1
; GENERAL INFORMATION:
; APPLICANT: GENFIT SA
; TITLE OF INVENTION: Compositions and Methods for the assay of AA4RP
; FILE REFERENCE: B0157W0
; CURRENT APPLICATION NUMBER: US/10/487,096
; CURRENT FILING DATE: 2004-02-19
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 95
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Description of artificial sequence : synthetic peptide AA4RP
US-10-487-096-2

Query Match          16.8%; Score 96.5; DB 17; Length 95;
Best Local Similarity 29.5%; Pred. No. 0.081;
Matches 23; Conservative 21; Mismatches 31; Indels 3; Gaps 2;

QY 24 DSGRDYVSQFEGSALGKQNLKLLDNWDSVTSTFSKLRQLGPTVQEFWDNLEKETGLR 83
| | | | | : : : | | : : : | | : : : | | : : : | | : : : | | : : :
Db 18 DKGR--VEGIHQQKWARE-PATLKDSLEQDLNNMKNFKLEKRLPLSGSEAPRLPQDPVGM 74

QY 84 QEMSKDLBVBKQVPYL 101
: : : : : | | : : : | | : : : | | : : : | | : : : | | : : :
Db 75 RQLQEEBVKARLQPYM 92

RESULT 10
US-10-142-238A-40
```

```
; Sequence 40, Application US/10142238A
; Publication No. US20030087819A1
; GENERAL INFORMATION:
; APPLICANT: Bielicki, John K.
; TITLE OF INVENTION: CYSTEINE-CONTAINING PEPTIDES HAVING OXIDANT PROPERTIES
; FILE REFERENCE: IB-1705
; CURRENT APPLICATION NUMBER: US/10/142,238A
; CURRENT FILING DATE: 2002-08-19
; PRIOR APPLICATION NUMBER: US 60/289,944
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 40
; LENGTH: 30
; TYPE: PRT
; ORGANISM: ARTIFICIAL SEQUENCE
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(30)
; OTHER INFORMATION: HUMAN GENETIC ORIGIN
US-10-142-238A-40

Query Match          16.3%; Score 94; DB 14; Length 30;
Best Local Similarity 95.0%; Pred. No. 0.034;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 42 LNLKLLDNWDSVTSTFSKLR 61
| | | | | | | | | | | | | | | | | | | |
Db 11 LPLKLLDNWDSVTSTFSKLR 30

RESULT 11
US-10-142-238A-41
; Sequence 41, Application US/10142238A
; Publication No. US20030087819A1
; GENERAL INFORMATION:
; APPLICANT: Bielicki, John K.
; TITLE OF INVENTION: CYSTEINE-CONTAINING PEPTIDES HAVING OXIDANT PROPERTIES
; FILE REFERENCE: IB-1705
; CURRENT APPLICATION NUMBER: US/10/142,238A
; CURRENT FILING DATE: 2002-08-19
; PRIOR APPLICATION NUMBER: US 60/289,944
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 41
; LENGTH: 30
; TYPE: PRT
; ORGANISM: ARTIFICIAL SEQUENCE
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(30)
; OTHER INFORMATION: HUMAN GENETIC ORIGIN
US-10-142-238A-41

Query Match          16.3%; Score 94; DB 14; Length 30;
Best Local Similarity 95.0%; Pred. No. 0.034;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 42 LNLKLLDNWDSVTSTFSKLR 61
| | | | | | | | | | | | | | | | | | | |
Db 11 LPLKLLDNWDSVTSTFSKLR 30

RESULT 12
US-10-142-238A-33
; Sequence 33, Application US/10142238A
; Publication No. US20030087819A1
; GENERAL INFORMATION:
; APPLICANT: Bielicki, John K.
; TITLE OF INVENTION: CYSTEINE-CONTAINING PEPTIDES HAVING OXIDANT PROPERTIES
; FILE REFERENCE: IB-1705
; CURRENT APPLICATION NUMBER: US/10/142,238A
```


/ CURRENT FILING DATE: 2002-08-19
/ PRIOR APPLICATION NUMBER: US 60/289,944
/ PRIOR FILING DATE: 2001-05-09
/ NUMBER OF SEQ ID NOS: 84
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 33
/ LENGTH: 18
/ TYPE: PRT
/ ORGANISM: ARTIFICIAL SEQUENCE
/ FEATURE:
/ NAME/KEY: PEPTIDE
/ LOCATION: (1)..(18)
/ OTHER INFORMATION: HUMAN GENETIC ORIGIN
US-10-142-238A-33

Query Match 16.0%; Score 92; DB 14; Length 18;
Best Local Similarity 100.0%; Pred. No. 0.028;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 LKLLDNWDSVTSTFSKLR 61
| | | | | | | | | | | | | | | | | |
DB 1 LKLLDNWDSVTSTFSKLR 18

RESULT 13
US-10-700-340-37
/ Sequence 37, Application US/10700340
/ Publication No. US20040203023A1
/ GENERAL INFORMATION:
/ APPLICANT: Herath, Herath Mudiyansele Athula Chandrasiri
/ TITLE OF INVENTION: PROTEINS, GENES AND THEIR USE FOR DIAGNOSIS AND TREATMENT OF BREAST
/ FILE REFERENCE: 2543-1-032
/ CURRENT APPLICATION NUMBER: US/10/700,340
/ PRIOR FILING DATE: 2003-11-03
/ PRIOR APPLICATION NUMBER: PCT/GB02/02022
/ PRIOR FILING DATE: 2002-05-02
/ PRIOR APPLICATION NUMBER: GB0110790.3
/ PRIOR FILING DATE: 2001-05-02
/ PRIOR APPLICATION NUMBER: GB0118385.4
/ PRIOR FILING DATE: 2001-07-27
/ PRIOR APPLICATION NUMBER: GB0119791.2
/ PRIOR FILING DATE: 2001-08-14
/ PRIOR APPLICATION NUMBER: GB0120045.0
/ PRIOR FILING DATE: 2001-08-16
/ PRIOR APPLICATION NUMBER: GB0128062.7
/ PRIOR FILING DATE: 2001-11-22
/ NUMBER OF SEQ ID NOS: 156
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 37
/ LENGTH: 16
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-700-340-37

Query Match 15.5%; Score 89; DB 17; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.048;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 EQLGPVTQBFWDNLEK 77
| | | | | | | | | | | | | | | | | |
DB 1 EQLGPVTQBFWDNLEK 16

RESULT 14
US-10-142-238A-39
/ Sequence 39, Application US/10142238A
/ Publication No. US20030087819A1
/ GENERAL INFORMATION:
/ APPLICANT: Bielicki, John K.
/ TITLE OF INVENTION: CYSTEINE-CONTAINING PEPTIDES HAVING OXIDANT PROPERTIES
/ FILE REFERENCE: IB-1705
/ CURRENT APPLICATION NUMBER: US/10/142,238A

/ CURRENT FILING DATE: 2002-08-19
/ PRIOR APPLICATION NUMBER: US 60/289,944
/ PRIOR FILING DATE: 2001-05-09
/ NUMBER OF SEQ ID NOS: 84
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 39
/ LENGTH: 30
/ TYPE: PRT
/ ORGANISM: ARTIFICIAL SEQUENCE
/ FEATURE:
/ NAME/KEY: PEPTIDE
/ LOCATION: (1)..(18)
/ OTHER INFORMATION: HUMAN GENETIC ORIGIN
US-10-142-238A-39

Query Match 15.5%; Score 89; DB 14; Length 30;
Best Local Similarity 90.0%; Pred. No. 0.1;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 42 LNLKLLDNWDSVTSTFSKLR 61
| | | | | | | | | | | | | | | | | |
DB 11 LPLKCLDNWDSVTSTFSKLR 30

RESULT 15
US-10-142-238A-43
/ Sequence 43, Application US/10142238A
/ Publication No. US20030087819A1
/ GENERAL INFORMATION:
/ APPLICANT: Bielicki, John K.
/ TITLE OF INVENTION: CYSTEINE-CONTAINING PEPTIDES HAVING OXIDANT PROPERTIES
/ FILE REFERENCE: IB-1705
/ CURRENT APPLICATION NUMBER: US/10/142,238A
/ CURRENT FILING DATE: 2002-08-19
/ PRIOR APPLICATION NUMBER: US 60/289,944
/ PRIOR FILING DATE: 2001-05-09
/ NUMBER OF SEQ ID NOS: 84
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 43
/ LENGTH: 30
/ TYPE: PRT
/ ORGANISM: ARTIFICIAL SEQUENCE
/ FEATURE:
/ NAME/KEY: PEPTIDE
/ LOCATION: (1)..(30)
/ OTHER INFORMATION: HUMAN GENETIC ORIGIN
US-10-142-238A-43

Query Match 15.5%; Score 89; DB 14; Length 30;
Best Local Similarity 90.0%; Pred. No. 0.1;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 42 LNLKLLDNWDSVTSTFSKLR 61
| | | | | | | | | | | | | | | | | |
DB 11 LPLKCLDNWDSVTSTFSKLR 30

Search completed: December 21, 2004, 12:24:19
Job time : 145 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2004, 12:22:02 ; Search time 38 Seconds
(without alignments)

275.990 Million cell updates/sec

Title: US-09-803-918A-2_COPY_25_133

Perfect score: 576

Sequence: 1 DEPPQSPWDRVKDLATVYVD.....LBEVKAKVQPYLDDFQKKWQ 109

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 44103

Minimum DB seq length: 0

Maximum DB seq length: 109

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_79:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	106	18.4	34	2 S67972	apolipoprotein AI
2	106	18.4	36	2 A56866	apolipoprotein A-I
3	101	17.5	20	2 A05313	apolipoprotein A-I
4	66.5	11.5	79	2 G86722	hypothetical prote
5	63	10.9	99	2 B46598	ski-related protei
6	63	10.9	104	2 A60094	neurofilament-M ho
7	63	10.9	107	1 H64112	virulence-associat
8	61.5	10.7	96	2 I68742	integral membrane
9	61	10.6	75	2 C90225	hypothetical prote
10	61	10.6	75	2 C90342	hypothetical prote
11	61	10.6	108	2 A59010	antifreeze protein
12	61	10.6	109	2 F71511	hypothetical prote
13	60.5	10.5	107	2 D95153	hypothetical prote
14	59.5	10.3	76	2 D82445	hypothetical prote
15	59.5	10.3	99	2 J00038	nonhistone chromos
16	59	10.2	102	2 D97844	hypothetical prote
17	59	10.2	109	2 F81669	conserved hypothet
18	58.5	10.2	76	2 JC5734	apolipoprotein A-I
19	58.5	10.2	79	2 AB3643	hypothetical cytos
20	58.5	10.2	85	2 C65120	hypothetical 10.0
21	58.5	10.2	93	2 AF3532	transposase BME110
22	58.5	10.2	93	2 AI3537	hypothetical prote
23	58	10.1	88	2 A72222	hypothetical prote
24	58	10.1	103	2 B69466	hypothetical prote
25	58	10.1	109	2 E34510	homeotic protein E
26	57.5	10.0	73	2 H81298	probable molybdopt
27	57.5	10.0	94	2 I55394	MHC HLA-DR7 DQ-bet
28	57.5	10.0	95	2 H69767	hypothetical prote
29	57.5	10.0	108	2 I72481	HLA DR-beta-I - hu

hypothetical prote
hypothetical prote
apolipoprotein C-I
M protein precurs
4-oxalocrotonate t
M-like protein enn
M-like protein enn
M-like protein enn
M-like protein enn
M-like protein enn
M-like protein enn
hypothetical prote
hypothetical prote
conserved domain p
transforming prote
homeotic protein E
hypothetical prote

ALIGNMENTS

RESULT 1

S67972

apolipoprotein AI - goose (fragment)

C;Species: Anser anser (domestic goose)

C;Date: 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004

C;Accession: S67972

R;Hermier, D.; Sellier, N.; Rousselot-Pailley, D.; Forgez, P.

Eur. J. Biochem. 234, 586-591, 1995

A;Title: Characterization of apolipoproteins B-100, AI and C from plasma lipoprotein in

A;Reference number: S67972; MUID:96128192; PMID:8536707

A;Accession: S67972

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-34 <HER>

A;Cross-references: UNIPROT:Q9PRR6

C;Superfamily: apolipoprotein A-I

Query Match 18.4%; Score 106; DB 2; Length 34;

Best Local Similarity 54.5%; Pred. No. 0.0093; 6; Indels 0; Gaps 0;

Matches 18; Conservative 9; Mismatches 6; Indels 0; Gaps 0;

QY 2 EPPQSPWDRVKDLATVYVDVLKDSGRDYVSQFE 34

Db 1 DEPPQELDLKDLVDVYLTQKASGKDYLAQFE 33

RESULT 2

A56866

apolipoprotein A-I - Japanese quail (fragment)

C;Species: Coturnix coturnix japonica (Japanese quail)

C;Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 09-Jul-2004

C;Accession: A56866

R;Oku, H.; Ishikawa, M.; Nagata, J.; Toda, T.; Chinen, I.

Biochim. Biophys. Acta 1167, 22-28, 1993

A;Title: Lipoprotein and apoprotein profile of Japanese quail.

A;Reference number: A56866; MUID:93213845; PMID:8461329

A;Accession: A56866

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-36 <OKU>

A;Cross-references: UNIPROT:P32918

A;Note: sequence extracted from NCBI backbone (NCBI:128831)

A;Note: this protein was found primarily as a 26K apoprotein

C;Superfamily: apolipoprotein A-I

Query Match 18.4%; Score 106; DB 2; Length 36;

Best Local Similarity 50.0%; Pred. No. 0.0099; 8; Indels 0; Gaps 0;

Matches 18; Conservative 10; Mismatches 8; Indels 0; Gaps 0;

QY 2 EPPQSPWDRVKDLATVYVDVLKDSGRDYVSQFE 37

Db 1 DDPQTPDLRDMLDVYLTQKASGKDAISQFE 36


```
A;Accession: F71511  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-109 <ARN>  
A:Cross-references: UNIPROT.O84472; GB:AEOO1J320; GB:AEOO1Z73; NID:G3328891; PIDN:AAC6806  
A:Experimental source: serotype D, strain UW-3/Cx  
C:Genetics:  
A:Gene: CT466  
C:Superfamily: conserved hypothetical protein CP0165
```

	Query Match	10.6%; Score 61;	DB 2;	Length 109;	
	Best Local Similarity	25.3%;	Pred. No.	2.le+02;	
	Matches	20;	Conservative	23;	
		Mismatches	28;	Indels	8;
	Gaps				3;

Qy	30	VSEFGSALCKQLKLLNDWDSVTSFFSKLRSLQGVPVTQEFWDNLKEKETEGLRQESKD	89
	:	: :: :	:: :
	Dd	23 LTQEASLSPEDL-IKVLRQQKLTLSCIEKHDI-----KKFRDSP---SLALPQEVQEE	74
	:	:: :	:: :
Qy	90	LEEVKAQPYPDLDFOKKW	108
	:	:	:: :
Db	75	LEGIRSVIQRILETDKNKY	93
	:	:	:: :

RESULT 13

D95153
hypothetical protein Spi323 [imported] - Streptococcus pneumoniae (strain TIGR4)
C:Species: Streptococcus pneumoniae
C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
C:Accession: D95153
R:Jettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
on, J.D.; Unayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A:Reference number: A95000; MUID:21357209; PMID:11463916
A:Accession: D95153
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-107 <KUR>
A:Cross-references: UNIPROT:Q97QA2; GB:A5005672; PIDN:AAK75421.1; PID:g14972804; GSPDB:G
A:Experimental source: strain TIGR4
C:Genetics:
A:Gene: Spi323

	Query Match	10.5%; Score 60.5; DB 2;	Length 107;
	Best Local Similarity	23.5%; Pred. No. 2.2e+02;	
	Matches	Conservative 17; Mismatches 35; Indels 23; Gaps 2;	
Qy	6 SPWDRVKDLATVYVDLKDGRDY-----VSQEGSALGQLNLKLLDNWDSTFESKL 60	: : : : :	: :
Dd	4 SVFTMQDIENAVTDIIKSVDNEITYTYKAQSQELEKLSYDEKSHEELVSIET----- 58	:	: :
Qy	61 REQLPVTQFWNDLNKEKETGLRQEMSKDLEEVAKVQ 98	: : :	:
Dd	59 -----NLEMQQNLIDEVNKTIKENDANIQ 83	: : :	:

RESULT 14
D82445
hypothetical protein VCA0551 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C/Species: Vibrio cholerae
C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C/Accession: D82445
R/Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
Chardon, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A/Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A/Reference number: A82035; MUID:20406833; PMID:10952301
A/Accession: D82445
A/Status: preliminary
A/Molecule type: DNA

A;Residues: 1-76 <HI>
A;Cross-references: UNIPROT:Q9KM38; GB:AE004386; GB:AE003853; NID:g9657957; PIDN:AAF964;
A;Experimental source: serogroup O1; strain N16961; biotype E1 Tor
C;Genetics:
A;Gene: VCA0551
A;Map position: 2

```
Query Match      10.3%; Score 59.5; DB 2; Length 76;
Best Local Similarity 22.0%; Pred. No. 1.8e+02;
Matches 11; Conservative 11; Mismatches 13; Indels 15; Gaps 1;
```

RESULT 15

JU0038 nonhistone chromosomal protein HMG - Tetrahymena pyriformis

C:Species: Tetrahymena pyriformis

C:Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 16-Aug-2004

C:Accession: JU0038

R:Hayashi, T.; Hayashi, H.; Iwai, K.

J. Biochem. 105, 577-581, 1989

A:Title: Tetrahymena HMG nonhistone chromosomal protein. Isolation and amino acid sequen

A:Reference number: JU0038; MUID:89340384; PMID:2760016

A:Accession: JU0038

A:Molecule type: Protein

A:Residues: 1-99 <HAY>

A:Cross-references: UNIPROT:P40625

A:Note: this HMG is similar to the central part of vertebrate HMG1

A:Note: Ser-16 and Thr-42 and/or Ser-43 are 6-7% phosphorylated

C:Genetics:

A:Genetic code: SGC5
C:Superfamily: HMG box homology
C:Keywords: chromosomal protein; DNA binding; nucleus; phosphoprotein
F:8-83/Domain: HMG box homology <HMG1>
F:16,43/Binding site: phosphate (Ser) (covalent) (partial) #status experimental
F:42/Binding site: phosphate(Thr) (covalent) (partial) #status experimental

Query Match	10.3%	Score 59.5;	DB 2;	Length 99;
Best Local Similarity	21.1%;	Pred. No. 2.5e+02;		
Matches	23;	Conservative 18;	Mismatches 32;	Indels 33; Gaps 4;
Qy	3	PPQSPMDRVKDLATVVDLKDGRDVSQFEGSALGQLNLKLLDNWDSVSTFSKLRLE	62	
Db	10	PPKRP-----LSAFF--LFQHNYPQVKNENPAKITELTSMIAEKWHT-----	53	
Qy	63	QLQGPVTQEFWDNLKEKETGLRQEMSKDLBEVKAQVQPYLDDFOKKW	108	
Db	54	-----EKKKKVEGLQO-----EAKAKYKDMOAAVEKKY	82	

Search completed: December 21, 2004, 12:34:46
Job time : 43 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2004, 12:19:19 ; Search time 39 Seconds
(without alignments)
185.350 Million cell updates/sec

Title: US-09-803-918A-2_COPY_25_133

Perfect score: 576

Sequence: 1 DEPPQSPWRVKDLATYYVD.....LEEKAKVQPYLDDFOKKWQ 109

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 317198

Minimum DB seq length: 0

Maximum DB seq length: 109

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/1/iaa/5A COMB.pep.*

2: /cgn2_6/ptodata/1/iaa/5B COMB.pep.*

3: /cgn2_6/ptodata/1/iaa/6A COMB.pep.*

4: /cgn2_6/ptodata/1/iaa/6B COMB.pep.*

5: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pep.*

6: /cgn2_6/ptodata/1/iaa/backfileei.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	143	24.8	44	2	US-08-292-870-2
2	133	23.1	64	2	US-08-292-870-1
3	114	19.8	25	2	US-08-292-870-4
4	83	14.4	32	2	US-08-292-870-3
5	81	14.1	16	1	US-07-959-946-5
6	81	14.1	16	1	US-08-333-577-5
7	81	14.1	16	5	PCT-US92-08634-5
8	67.5	11.7	87	1	US-08-685-764-4
9	67	11.6	105	4	US-09-513-999C-6681
10	66.5	11.5	105	1	US-08-241-853-11
11	66.5	11.5	105	2	US-08-850-917-11
12	65.5	11.4	99	4	US-09-147-875A-16
13	63.5	11.0	50	3	US-09-695-458-9
14	63	10.9	103	4	US-09-270-767-57235
15	63	10.9	106	4	US-09-621-976-4439
16	62	10.8	65	2	US-08-867-087B-34
17	61.5	10.7	98	2	US-08-479-078-7
18	61.5	10.7	99	2	US-08-710-749-10
19	61.5	10.7	99	4	US-09-147-875A-11
20	61	10.6	66	2	US-08-867-087B-70
21	61	10.6	100	4	US-09-147-875A-12
22	60.5	10.5	37	3	US-09-695-458-15
23	60	10.4	80	3	US-09-183-861-61
24	60	10.4	80	3	US-08-022-765-61
25	60	10.4	80	4	US-09-551-974A-61
26	60	10.4	80	4	US-09-565-501A-61
27	60	10.4	80	4	US-09-639-206A-61

ALIGNMENTS

RESULT 1

US-08-292-870-2
; Sequence 2, Application US/08292870
; Patent No. 5814467

; GENERAL INFORMATION:

; APPLICANT: Curtiss, Linda K

; APPLICANT: Banka, Carole L

; APPLICANT: Bonnet, David J

; APPLICANT: Smith, Richard S

; TITLE OF INVENTION: APO AI POLYPEPTIDES, DIAGNOSTIC METHODS

; TITLE OF INVENTION: AND SYSTEMS FOR QUANTIFYING APO AI, AND THERAPEUTIC

; TITLE OF INVENTION: METHODS

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: The Scripps Research Institute, Office of

; ADDRESSEE: Patent Counsel

; STREET: 10666 No. 5814467th Torrey Pines Road., TPC 8

; CITY: La Jolla

; STATE: California

; COUNTRY: US

; ZIP: 92037

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/292,870

; FILING DATE: 17-AUG-1994

; CLASSIFICATION: 436

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/534,761

; FILING DATE: 07-JUN-1990

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/711,333

; FILING DATE: 06-JUN-1991

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/US 91/04038

; FILING DATE: 07-JUN-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: Fitting, Thomas

; REGISTRATION NUMBER: 34,163

; REFERENCE/DOCKET NUMBER: TSRI 210.1 D1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 619-554-2937

; TELEFAX: 619-554-6312

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 44 amino acids

; TYPE: amino acid

Sequence 61, Appl
Sequence 12, Appl
Sequence 13, Appl
Sequence 28, Appl
Sequence 28, Appl
Sequence 15, Appl
Sequence 33, Appl
Sequence 158, App
Sequence 26113, A
Sequence 8, Appl
Sequence 8, Appl
Sequence 17, Appl
Sequence 47, Appl
Sequence 6615, Ap
Sequence 11, Appl
Sequence 3, Appl
Sequence 25, Appl


```
/
/ GENERAL INFORMATION:
/ APPLICANT: Curtiss, Linda K
/ APPLICANT: Banks, Carole L
/ APPLICANT: Bonnet, David J
/ APPLICANT: Smith, Richard S
/ TITLE OF INVENTION: APO AI POLYPEPTIDES, DIAGNOSTIC METHODS
/ TITLE OF INVENTION: AND SYSTEMS FOR QUANTIFYING APO AI, AND THERAPEUTIC
/ TITLE OF INVENTION: METHODS
/ NUMBER OF SEQUENCES: 4
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: The Scripps Research Institute, Office of
/ ADDRESSEE: Patent Counsel
/ STREET: 10666 No. 581467th Torrey Pines Road., TPC 8
/ CITY: La Jolla
/ STATE: California
/ COUNTRY: US
/ ZIP: 92037
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/292,870
/ FILING DATE: 17-AUG-1994
/ CLASSIFICATION: 436
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/534,761
/ FILING DATE: 07-JUN-1990
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/711,333
/ FILING DATE: 06-JUN-1991
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: PCT/US 91/04038
/ FILING DATE: 07-JUN-1991
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Fitting, Thomas
/ REGISTRATION NUMBER: 34,163
/ REFERENCE/DOCKET NUMBER: TSRI 210.1 D1
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 619-554-2937
/ TELEFAX: 619-554-6312
/ INFORMATION FOR SEQ ID NO: 3:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 32 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ FRAGMENT TYPE: internal
/ FEATURE:
/ NAME/KEY: Region
/ LOCATION: 11
/ OTHER INFORMATION: /note= "Xaa can be either E (Glu)
/
/ US-08-292-870-3
/
/ Query Match 14.4%; Score 83; DB 2; Length 32;
/ Best Local Similarity 93.8%; Pred. No. 0.071;
/ Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
/
/ QY 94 KAKVQPYLDDFQKKWQ 109
/ Db 1 KAKVQPYLDDXQKKWQ 16
/
/ RESULT 5
/ US-07-959-946-5
/ Sequence 5, Application US/07959946
/ Patent No. 5408038
/ GENERAL INFORMATION:
/ APPLICANT: Smith, Richard K.
/ APPLICANT: Koduri, Raju
/ APPLICANT: Young, Stephen G.
/
/ Query Match 14.4%; Score 81; DB 1; Length 16;
/ Best Local Similarity 100.0%; Pred. No. 0.048;
/ Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
/
/ QY 96 KVQPYLDDFQKKWQ 109
/ Db 1 KVQPYLDDFQKKWQ 14
/
/ RESULT 6
/ US-08-333-577-5
/ Sequence 5, Application US/08333577
/ Patent No. 5786206
/ GENERAL INFORMATION:
/ APPLICANT: Smith, Richard K.
/ APPLICANT: Koduri, Raju
/ APPLICANT: Young, Stephen G.
/ APPLICANT: Witztum, Joseph L.
/ APPLICANT: Curtiss, Linda K.
/ TITLE OF INVENTION: Lipoprotein Assays Using Antibodies to a
/ TITLE OF INVENTION: Pan Native Epitope and Recombinant Antigens
/ NUMBER OF SEQUENCES: 20
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Dressler, Goldsmith, Shore, Sutker &
/ ADDRESSEE: Milnamow, Ltd.
/ STREET: 180 No. 5786206th Stetson, Suite 4700
/ CITY: Chicago
/ STATE: Illinois
/ COUNTRY: USA
/ ZIP: 60601
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
```

```
/
/ APPLICANT: Witztum, Joseph L.
/ APPLICANT: Curtiss, Linda K.
/ TITLE OF INVENTION: Lipoprotein Assays Using Antibodies to a
/ TITLE OF INVENTION: Pan Native Epitope and Recombinant Antigens
/ NUMBER OF SEQUENCES: 20
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Dressler, Goldsmith, Shore, Sutker &
/ ADDRESSEE: Milnamow, Ltd.
/ STREET: 180 No. 5408038th Stetson, Suite 4700
/ CITY: Chicago
/ STATE: Illinois
/ COUNTRY: USA
/ ZIP: 60601
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/07/959,946
/ FILING DATE: 19921008
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/901,706
/ FILING DATE: 18-JUN-1992
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Gamson, Edward P.
/ REGISTRATION NUMBER: 29,381
/ REFERENCE/DOCKET NUMBER:
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (312)616-5400
/ TELEFAX: (312)616-5460
/ INFORMATION FOR SEQ ID NO: 5:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 16 amino acids
/ TYPE: AMINO ACID
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/
/ US-07-959-946-5
/
/ Query Match 14.1%; Score 81; DB 1; Length 16;
/ Best Local Similarity 100.0%; Pred. No. 0.048;
/ Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
/
/ QY 96 KVQPYLDDFQKKWQ 109
/ Db 1 KVQPYLDDFQKKWQ 14
/
/ RESULT 6
/ US-08-333-577-5
/ Sequence 5, Application US/08333577
/ Patent No. 5786206
/ GENERAL INFORMATION:
/ APPLICANT: Smith, Richard K.
/ APPLICANT: Koduri, Raju
/ APPLICANT: Young, Stephen G.
/ APPLICANT: Witztum, Joseph L.
/ APPLICANT: Curtiss, Linda K.
/ TITLE OF INVENTION: Lipoprotein Assays Using Antibodies to a
/ TITLE OF INVENTION: Pan Native Epitope and Recombinant Antigens
/ NUMBER OF SEQUENCES: 20
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Dressler, Goldsmith, Shore, Sutker &
/ ADDRESSEE: Milnamow, Ltd.
/ STREET: 180 No. 5786206th Stetson, Suite 4700
/ CITY: Chicago
/ STATE: Illinois
/ COUNTRY: USA
/ ZIP: 60601
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
```

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/333,577
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Gamson, Edward P.
; REGISTRATION NUMBER: 29,381
; REFERENCE/DOCKET NUMBER: SCRF 234.0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312)616-5400
; TELEFAX: (312)616-5460
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-333-577-5

Query Match      14.1%; Score 81; DB 1; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.048;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      96 KVQPYLDDFQKKWQ 109
Db      1 KVQPYLDDFQKKWQ 14

RESULT 7
PCT-US92-08634-5
; Sequence 5, Application PC/TUS9208634
; GENERAL INFORMATION:
; APPLICANT: Smith, Richard K.
; APPLICANT: Koduri, Raju
; APPLICANT: Young, Stephen G.
; APPLICANT: Witzum, Joseph L.
; APPLICANT: Curtiss, Linda K.
; TITLE OF INVENTION: Lipoprotein Assays Using Antibodies to a
; TITLE OF INVENTION: Pan Native Epitope and Recombinant Antigens
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dressler, Goldsmith, Shore, Sukter &
; ADDRESSEE: Milnamow, Ltd.
; STREET: 180 North Stetson, Suite 4700
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/08634
; FILING DATE: 19921009
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/901,706
; FILING DATE: 18-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Gamson, Edward P.
; REGISTRATION NUMBER: 29,381
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312)616-5400
; TELEFAX: (312)616-5460
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: AMINO ACID
```

```
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US92-08634-5

Query Match      14.1%; Score 81; DB 5; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.048;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      96 KVQPYLDDFQKKWQ 109
Db      1 KVQPYLDDFQKKWQ 14

RESULT 8
US-08-685-764-4
; Sequence 4, Application US/08685764
; Patent No. 5800982
; GENERAL INFORMATION:
; APPLICANT: HASEGAWA, AKIRA
; APPLICANT: MAKI, NOBORU
; APPLICANT: YAGI, SHINTARO
; APPLICANT: KASHIWAKUMA, TOMIKO
; TITLE OF INVENTION: ANTIGENIC PEPTIDES FOR GROUPING
; TITLE OF INVENTION: HEPATITIS C VIRUS, KIT COMPRISING THE SAME AND
; TITLE OF INVENTION: METHODS FOR ITS GROUPING USING THE SAME
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
; ADDRESSEE: CUSHMAN
; STREET: 130 WATER STREET
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/685,764
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/092,192
; FILING DATE: 15-JUL-1993
; APPLICATION NUMBER: JP 212061/92
; FILING DATE: 16-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 316634/92
; FILING DATE: 30-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 316635/92
; FILING DATE: 30-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 104754/93
; FILING DATE: 30-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: BUCKLEY, LINDA M.
; REGISTRATION NUMBER: 31003
; REFERENCE/DOCKET NUMBER: 42822
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 523-3400
; TELEFAX: (617) 523-6440
; TELEX: 200291 STRE UR
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 87 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-685-764-4
```

Query Match 11.7%; Score 67.5; DB 1; Length 87;
Best Local Similarity 24.2%; Pred. No. 9.7;
Matches 23; Conservative 22; Mismatches 39; Indels 11; Gaps 3;

QY 13 DLATVVVDVLKDSGRDVVSQFEGSALGKQLNLKLLDNWDSVTSTFSKLRQLGPTQEFW 72
DB 1 EFAFGCVSII--GRLLHINQRAVVPDKVELYFAFDEMEECASRAALIEB-----GQRTA 52
QY 73 DNLKEKETGLQRQMSKDLEBVKAKVQ---PYLDDF 104
DB 53 EMLKSKIQGLLQQAASKQAQDIKPAVQTSMPKVEQF 87

RESULT 9
US-09-513-999C-6681
; Sequence 6681, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 6681
; LENGTH: 105
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-513-999C-6681

Query Match 11.6%; Score 67; DB 4; Length 105;
Best Local Similarity 28.4%; Pred. No. 14;
Matches 19; Conservative 16; Mismatches 28; Indels 4; Gaps 2;

QY 39 GKQLNLKLLD-NWDSVTSTFSKLRQLGPTQEFWNLEKETEG---LRQMSKDLEBVK 94
DB 7 GREQQDIMNKYQQLSRLDEILSIKAKETEBIKQLTEGQIAANEALKKQLEGI 66
QY 95 AKVQPYL 101
DB 67 SGLQLEYL 73

RESULT 10
US-08-241-853-11
; Sequence 11, Application US/08241853
; Patent No. 5693488
; GENERAL INFORMATION:
; APPLICANT: Fang, Kathy S.
; APPLICANT: Hanafusa, Hidesaburo
; TITLE OF INVENTION: A TRANSMEMBRANE TYROSINE PHOSPHATASE
; TITLE OF INVENTION: AND METHODS OF USE THEREOF
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/241,853

; FILING DATE: 12-MAY-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-078
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-241-853-11

Query Match 11.5%; Score 66.5; DB 1; Length 105;
Best Local Similarity 27.7%; Pred. No. 15;
Matches 26; Conservative 11; Mismatches 28; Indels 29; Gaps 4;
QY 8 WDRVKDLATVYVDVLKD-----SGRDYVSQFEGS--ALGK 40
DB 12 WDQLADLAAPCKKNLQDAENPFQODADDLKAWLQDAHRLLSGED-VQDDEGATRALGK 70
QY 41 QLNKLLDNWDSVTSTFSKLRQLGPTQEFWDN 74
DB 71 K-HKDFLEBESRGVMEKLEQQAQGFPEPRDS 103

RESULT 11
US-08-850-917-11
; Sequence 11, Application US/08850917
; Patent No. 5854045
; GENERAL INFORMATION:
; APPLICANT: Fang, Kathy S.
; APPLICANT: Hanafusa, Hidesaburo
; TITLE OF INVENTION: A TRANSMEMBRANE TYROSINE PHOSPHATASE
; TITLE OF INVENTION: AND METHODS OF USE THEREOF
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/850,917
; FILING DATE: 02-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/241,853
; FILING DATE: 12-MAY-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-078
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:

```

; LENGTH: 105 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-950-917-11

Query Match      11.5%; Score 66.5; DB 2; Length 105;
Best Local Similarity 27.7%; Pred. No. 15;
Matches 26; Conservative 11; Mismatches 28; Indels 29; Gaps 4;

Qy      8 WDRVKDIAIVYDVLDK-----SQRDVVSQFEGS--ALGK 40
      |||:|||||:|:|
Db     12 WDQQLKDAAFCKKNLQDAENFFQFGDADDLKAWLQDAHRLLSGED-VGQDEGATRALGK 70
      |||:|||||:|:|

Qy     41 QLNKLKLDNWDVSTSTFSKLRQQLGPVTQBFWDN 74
      |||:|:|:|:|:|:|:|
Db     71 K-HKDFLEELLESRGVMEKLEQQQAQGFPEFRDS 103
      |||:|:|:|:|:|:|:|

RESULT 12
US-09-147-875A-16
; Sequence 16, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-147-875A-16

Query Match      11.4%; Score 65.5; DB 4; Length 99;
Best Local Similarity 24.2%; Pred. No. 18;
Matches 24; Conservative 19; Mismatches 37; Indels 19; Gaps 3;

Qy     20 DVLKDSGRDVVSQFEGSALGQLNLK-----LLDNWDSVTSTFSKLRQQL----- 64
      ::|:|:|:|:|:|:|:|
Db     3 EIDESSEDYLVKEGERAPLQSKLDTTKAKLSKLELSDKIDELDAETAKLEVQLKDAEGN 62
      ::|:|:|:|:|:|:|:|

Qy     65 GPVTQBFWDNLEKTEGLRQEMSKDLSEVAKVQPYLDD 103
      |||:|:|:|:|:|:|:|
Db     63 NNVEAYPFKEGLEKTT---AEKKAEELEKAEADLKKAYDE 97
      |||:|:|:|:|:|:|:|

RESULT 13
US-09-695-458-9
; Sequence 9, Application US/09695458
; Patent No. 6380361
; GENERAL INFORMATION:
; APPLICANT: Conklin, Darrell C.
; APPLICANT: Taft, David W.
; TITLE OF INVENTION: Educational Kit and Method Containing No. 6380361el Alpha Helical
; TITLE OF INVENTION: Protein-34
; FILE REFERENCE: 99-78
; CURRENT APPLICATION NUMBER: US/09/695,458
; CURRENT FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/162,623
; PRIOR FILING DATE: 1999-10-29
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 50
; TYPE: PRT
; ORGANISM: Mus musculus

```

Search completed: December 21, 2004, 12:33:58
Job time : 43 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2004, 12:05:04 ; Search time 38 Seconds
(without alignments)
98.749 Million cell updates/sec

Title: US-09-803-918A-2_COPY_75_113
Perfect score: 201
Sequence: 1 DSVTSTFSKRLQGLGVTQEFWDNLEKETEGLRQEMSKD 39
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 9448

Minimum DB seq length: 0
Maximum DB seq length: 41

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	40	19.9	33	2	S52151	hypothetical prote
2	37	18.4	40	2	H81592	hypothetical prote
3	36	17.9	34	2	G84147	hypothetical prote
4	36	17.9	35	2	AH1838	cytochrome c6 limp
5	34	16.9	19	2	E60977	14-3-3 protein hom
6	34	16.9	30	2	S74192	crotoxin inhibitor
7	34	16.9	40	2	T07516	hypothetical prote
8	33	16.4	39	2	G82613	hypothetical prote
9	32.5	16.2	27	2	PQ0844	DNA-binding protei
10	32	15.9	31	2	PS0350	myohemerythrin iso
11	32	15.9	34	2	C56635	tubulin alpha chai
12	32	15.9	35	2	A54257	deoxynucleoside ki
13	32	15.9	35	2	B82012	hypothetical prote
14	32	15.9	39	2	S70798	Ivi protein I - Vi
15	32	15.9	40	2	C95169	hypothetical prote
16	31	15.4	33	2	E82856	hypothetical prote
17	31	15.4	33	2	B56635	tubulin alpha chai
18	31	15.4	37	2	S26954	peptide YY-related
19	31	15.4	40	2	A59005	thymosin beta - se
20	30.5	15.2	29	2	S17432	H4-transporting tw
21	30	14.9	15	2	PA0091	methionine adenosy
22	30	14.9	27	1	SECH	secretin - chicken
23	30	14.9	29	2	C61384	tracheal mucin gly
24	30	14.9	30	2	H72312	hypothetical prote
25	30	14.9	35	2	D81622	hypothetical prote
26	30	14.9	35	2	S53728	fructosyllysine-9p
27	30	14.9	36	2	PT0430	leucyl aminopeptid
28	30	14.9	36	2	A60343	transforming prote
29	30	14.9	36	2	C64039	hypothetical prote

30	30	14.9	40	2	S01182	hypothetical prote
31	30	14.9	40	2	C72398	hypothetical prote
32	29.5	14.7	27	2	H81080	hypothetical prote
33	29.5	14.7	28	2	A60303	vasoactive intesti
34	29.5	14.7	37	2	A97917	conserved hypothet
35	29	14.4	16	2	S28213	glutathione transf
36	29	14.4	20	2	S29099	glutathione transf
37	29	14.4	20	2	S30381	glutathione transf
38	29	14.4	24	2	S17926	glutathione transf
39	29	14.4	35	2	JH0639	GTP-binding protei
40	29	14.4	36	2	S27054	neuropeptide Y - A
41	29	14.4	39	2	S65949	hypothetical prote
42	29	14.4	39	2	C55995	prostaglandin E2 r
43	29	14.4	39	2	PQ0011	tubulin beta chain
44	29	14.4	40	2	I39944	regulatory extrace
45	29	14.4	41	2	T48342	hypothetical prote

ALIGNMENTS

RESULT 1

S52151 hypothetical protein - phage 2C (fragment)

C:Species: phage 2C

A:Note: host Bacillus subtilis

C:Date: 19-Mar-1997 #sequence_revision 18-Jul-1997 #text_change 18-Jul-1997

C:Accession: S52151

R:Hoet, P.

submitted to the EMBL Data Library, December 1993

A:Description: Isolation and characterization of Bacillus subtilis phage 2C transcript

A:Reference number: S52151

A:Accession: S52151

A:Molecule type: DNA

A:Residues: 1-33 <HOE>

A:Cross-references: EMBL:X75889

C:Genetics:

A:Start codon: GTG

Query Match 19.9%; Score 40; DB 2; Length 33;
Best Local Similarity 47.6%; Pred. No. 1.7e+02;
Matches 10; Conservative 2; Mismatches 5; Indels 4; Gaps 1;

QY 15 GPVTQEFWDNLE----KETEG 31
Db 6 GAVTNEFKDRLQVFNKEVQG 26

RESULT 2

H81592

hypothetical protein CP0296 [imported] - Chlamydomophila pneumoniae (strain AR39)

C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae

C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004

C:Accession: H81592

R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heideberg, J.F.; White, O.; Hickey,

, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,

Nucleic Acids Res. 28, 1397-1406, 2000

A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.

A:Reference number: AB1500; MUID:20150255; PMID:10684935

A:Accession: H81592

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-40 <REA>

A:Cross-references: UNIPROT:Q9K2A3; GB:AE002191; GB:AE002161; NID:g7189216; PIDN:AAF381E

A:Experimental source: strain AR39, HL cells

C:Genetics:

A:Gene: CP0296

Query Match 18.4%; Score 37; DB 2; Length 40;
Best Local Similarity 30.0%; Pred. No. 4.9e+02;
Matches 6; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

QY 1 DSVTSTFSKRLQGLGVTQEQE 20

G82613
Hypothetical protein XF1988 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: G82613
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing
Nature 406, 151-157, 2000
A>Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A>Note: for a complete list of authors see reference number A59328 below
A:Accession: G82613
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-39 <SIM>
A:Cross-references: UNIPROT:Q9PBZ7; GB:AE004018; GB:AE003849; NID:g9107093; PIDN:AAF8479
A:Experimental source: strain 9a5c
R:Simpton, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A. Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H. as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E. A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.; P.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A. Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir M.; Tsuchiko, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z A:Reference number: A59328
A:Contents: annotation
A:Genetics:
A:Gene: XF1988
Query Match 16.4%; Score 33; DB 2; Length 39;
Best Local Similarity 38.9%; Pred. No. 1.5e+03;
Matches 7; Conservative 3; Mismatches 8; Indels 0; Gaps 0;
QY 12 EQLGPTQBFWDNLEKET 29
DB 12 EQKAVNDPFWISLKV 29
RESULT 9
PQ0844
DNA-binding protein AcBBP1 - Azorhizobium caulinodans (fragment)
C:Species: Azorhizobium caulinodans
C>Date: 03-May-1994 #sequence_revision 07-Oct-1994 #text_change 09-Jul-2004
C:Accession: PQ0844
R:Welter, P.; Metz, B.; Felix, G.; Palme, K.; Szczyglowski, K.; de Bruijn, F.J.
Plant Physiol. 102, 1095-1107, 1993
A>Title: Interaction of a rhizobial DNA-binding protein with the promoter region of a pl A:Reference number: PQ0844; MUID:94105338; PMID:8278541
A:Accession: PQ0844
A:Molecule type: protein
A:Residues: 1-27 <WEL>
A:Cross-references: UNIPROT:Q7MLA8
A:Experimental source: strain ORS571
Query Match 16.2%; Score 32.5; DB 2; Length 27;
Best Local Similarity 44.4%; Pred. No. 1.1e+03;
Matches 8; Conservative 5; Mismatches 4; Indels 1; Gaps 1;
QY 3 VTSTFSKLREQLGPTQEE 20
DB 7 VGRNFARLQKGLTQE 23
RESULT 10
PS0350
myohemerythrin isoform II - sipunculid (Phascolopsis gouldii) (fragment)
C:Species: Phascolopsis gouldii
C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
C:Accession: PS0350

R:Long, R.C.; Zhang, J.H.; Kurtz Jr., D.M.; Negri, A.; Tedeschi, G.; Bonomi, F.
submitted to JIPID, April 1992
A:Reference number: JS0678
A:Accession: PS0350
A:Molecule type: protein
A:Residues: 1-31 <ION>
A:Cross-references: UNIPROT:P27687
A:Comment: This myohemerythrin is a monomeric oxygen-binding protein found in the retra C:Superfamily: hemerythrin
C:Keywords: oxygen carrier
Query Match 15.9%; Score 32; DB 2; Length 31;
Best Local Similarity 50.0%; Pred. No. 1.5e+03;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 21 FWDNLEKETEGRLQ 34
DB 17 FYDLLDDHKGGLFQ 30
RESULT 11
CS6635
tubulin alpha chain, brain-specific isotype (clone pTUB6) - chum salmon (fragment)
C:Species: Oncorhynchus keta (chum salmon)
C>Date: 11-Aug-1995 #sequence_revision 11-Aug-1995 #text_change 10-Jul-1998
C:Accession: CS6635
R:Coe, I.R.; Munro, R.; Sherwood, N.M.
DNA Seq. 3, 257-262, 1992
A>Title: Isolation of different brain-specific isotypes of alpha-tubulins from chum sal A:Reference number: A56635; MUID:93208376; PMID:1296820
A:Contents: brain
A:Accession: CS6635
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-34 <COE>
A>Note: sequence extracted from NCBI backbone (NCBIN:128392, NCBI:128391)
C:Superfamily: tubulin
F:28/Binding site: polyglutamate (Glu) (covalent) #status predicted
F:33-34/Cleavage site: Glu-Tyr (tubulin-specific carboxypeptidase) #status predicted
F:33-34/Cross-link: peptide (Glu-Tyr) (by tubulin-tyrosine ligase) #status predicted
Query Match 15.9%; Score 32; DB 2; Length 34;
Best Local Similarity 28.6%; Pred. No. 1.6e+03;
Matches 8; Conservative 8; Mismatches 8; Indels 4; Gaps 1;
QY 7 FSKLEQLGPTQEE----WDNLEKETE 30
DB 1 FSEAREDMALEKDYEEVGVDSIEGEGE 28
RESULT 12
AS4257
deoxynucleoside kinase complex I S-component - Lactobacillus acidophilus (fragment)
C:Species: Lactobacillus acidophilus
C>Date: 06-Oct-1994 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C:Accession: AS4257
R:Ikeda, S.; Ma, G.T.; Ives, D.H.
Biochemistry 33, 5328-5334, 1994
A>Title: Heterodimeric deoxynucleoside kinases of Lactobacillus acidophilus R-26: functi A:Reference number: AS4257; MUID:94227067; PMID:8172906
A:Accession: AS4257
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-35 <IKE>
A:Cross-references: UNIPROT:Q9R4Z3
A:Experimental source: R-26
A>Note: sequence extracted from NCBI backbone (NCBI:146748)
C:Superfamily: Lactobacillus acidophilus deoxyadenosine kinase
Query Match 15.9%; Score 32; DB 2; Length 35;
Best Local Similarity 39.1%; Pred. No. 1.7e+03;
Matches 9; Conservative 5; Mismatches 7; Indels 2; Gaps 1;

Qy	4	T	S	F	S	K	L	R	E	Q	L	G	P	V	T	Q	E	F	W	D	N	L	E	26	
	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	
db	13	S	S	L	S	L	L	A	E	X	L	G	--	T	O	A	F	Y	E	G	V	D			33

RESULT 13

B82012
 hypothetical protein NMA0180 [imported] - *Neisseria meningitidis* (strain Z2491 serogroup
 C)Species: *Neisseria meningitidis*
 C/Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
 C/Accession: B82012
 R/Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
 ; Holroyd, S.; Jagsels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
 Nature 404, 502-506, 2000
 A>Title: Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491.
 A/Reference number: A81775; MUID:20225558; PMID:10761919
 A/Accession: B82012
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-35 <PAR>
 A/Cross-references: UNIPROT:Q9JWX5; GB:AL162752; GB:AL157959; NID:G7378778; PIDN:CAB8349
 A/Experimental source: serogroup A, strain Z2491
 C/Genetics:
 A/Gene: NMA0180

Query Match 15.9%; Score 32; DB 2; Length 35;
Best Local Similarity 54.5%; Pred. No. 1.7e+03;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 9 KLREQLGPVTQ 19
Db 14 KYLEQFGPVS 24

RESULT 14

S70798
Ivi protein I - *Vibrio cholerae* (fragment)
N:Alternate names: *cysI* protein homolog
C:Species: *Vibrio cholerae*
C:Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 09-Jul-2004
C:Accession: S70798
R:Camilli, A.; Mekalanos, J.J.
Mol. Microbiol. 18, 671-683, 1995
A:Title: Use of recombinase gene fusions to identify *Vibrio cholerae* genes induced during
A:Reference number: S70798; MUID:96414463; PMID:8817490
A:Accession: S70798
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-39 <CAM>
A:Cross-references: UNIPROT:Q56599; EMBL:U25709; NID:g1165175; PIDN:AAC43551.1; PID:g116
C:Superfamily: sulfite reductase

Query Match 15.9%; Score 32; DB 2; Length 39;
Best Local Similarity 38.5%; Pred. No. 1.9e+03;
Matches 10; Conservative 6; Mismatches 8; Indels 2; Gaps 1;

QY 10 LREQLPVTQFEWDN--LEKETEGRL 33
: : | | | : | | | : : | |
pb 9 VOEVLGEVLGPWSDNERLKRSHFLR 34

RESULT 15

C59169
 hypothetical protein SPI454 [imported] - Streptococcus pneumoniae (strain TIGR4)
 C:Species: Streptococcus pneumoniae
 C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
 C:Accession: C95169
 R:fetzelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
 on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,
 n, T.; Hickey, E.K.; Holt, I.E.
 Science 293, 498-506, 2001
 A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
 A.; Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2004, 12:02:34 ; Search time 189 Seconds
(without alignments)
118.728 Million cell updates/sec

Title: US-09-803-918A-2_COPY_75_113

Perfect score: 201

Sequence: 1 DSVTSTFSKLRQLGPTQEFWDNLEKTEGLRQEMSKD 39

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 56135

Minimum DB seq length: 0

Maximum DB seq length: 41

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_02.*

1: uniprot_sprot.*

2: uniprot_trembl.*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	51.5	25.6	37	2 Q6KGF3	Q6kfg3 bacterioph
2	45.5	22.6	39	2 Q8WTG3	Q8wtg3 pepais sp.
3	45.5	22.6	39	2 Q8WT10	Q8wt10 cephus sp.
4	45	22.4	32	2 Q73DP3	Q73dp3 bacillus ce
5	45	22.4	32	2 AAS39602	Aas39602 bacillus
6	43	21.4	41	2 Q13587	Q13587 homo sapien
7	41	20.4	32	2 Q53103	Q53103 rhodobacter
8	40	19.9	40	2 Q62886	Q62886 canis famil
9	39.5	19.7	29	2 P97599	P97599 rattus norv
10	39.5	19.7	36	2 Q45208	Q45208 acanthokara
11	39	19.4	27	2 Q9MHZ4	Q9mhz4 drosophila
12	39	19.4	27	2 Q9MHZ6	Q9mhz6 drosophila
13	39	19.4	27	2 Q9MHZ8	Q9mhz8 drosophila
14	39	19.4	27	2 Q9MI20	Q9mi20 drosophila
15	39	19.4	27	2 Q9MI26	Q9mi26 drosophila
16	39	19.4	27	2 Q9MI28	Q9mi28 drosophila
17	39	19.4	27	2 Q9MI30	Q9mi30 drosophila
18	39	19.4	34	2 Q8LVV4	Q8lvv4 ceratitidis r
19	39	19.4	34	2 Q8LVV5	Q8lvv5 bactrocera
20	39	19.4	34	2 Q8LVV6	Q8lvv6 bactrocera
21	39	19.4	34	2 Q8LYC8	Q8lyc8 ceratitidis r
22	39	19.4	34	2 Q8LYD4	Q8lyd4 bactrocera
23	39	19.4	34	2 Q8LYD8	Q8lyd8 bactrocera
24	39	19.4	34	2 Q8LYD9	Q8lyd9 bactrocera
25	39	19.4	34	2 Q8LYE0	Q8lye0 bactrocera
26	39	19.4	34	2 Q8LYE2	Q8lye2 bactrocera
27	39	19.4	34	2 Q8LYE4	Q8lye4 bactrocera
28	39	19.4	34	2 Q8LYE6	Q8lye6 anastrepha
29	39	19.4	34	2 Q8LYE8	Q8lye8 anastrepha
30	38	18.9	26	2 Q7R9A2	Q7r9a2 plasmodium
31	38	18.9	27	2 Q9MI00	Q9mi00 drosophila

32	38	18.9	27	2 Q9MI02	Q9mi02 drosophila
33	38	18.9	27	2 Q9MI04	Q9mi04 drosophila
34	38	18.9	27	2 Q9MI06	Q9mi06 bactrocera
35	38	18.9	27	2 Q9MI14	Q9mi14 drosophila
36	38	18.9	27	2 Q9MI22	Q9mi22 drosophila
37	38	18.9	27	2 Q9MI24	Q9mi24 drosophila
38	38	18.9	30	2 Q9K532	Q9k532 listeria mo
39	38	18.9	34	2 Q8LVV3	Q8lvv3 ceratitidis c
40	38	18.9	34	2 Q8LYB7	Q8lyb7 toxotrypana
41	38	18.9	34	2 Q8LYB9	Q8lyb9 rhagoletis
42	38	18.9	34	2 Q8LYC5	Q8lyc5 dacus demme
43	38	18.9	34	2 Q8LYC7	Q8lyc7 dacus cilia
44	38	18.9	34	2 Q8LYD2	Q8lyd2 ceratitidis c
45	38	18.9	34	2 Q8LYD6	Q8lyd6 bactrocera

ALIGNMENTS

RESULT 1

Q6KGF3 PRELIMINARY; PRT; 37 AA.
AC Q6KGF3;
DT 05-JUL-2004 (Tremblrel. 27, Created)
DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)
DE Hypothetical protein.
DE Hypothetical protein.
OS Bacteriophage Felix 01.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales.
OX NCBI_TaxID=77775;
RN [1]
RP SEQUENCE FROM N.A.
RA Srianganathan N., Whichard J.M., Pierson F.W., Kapur V., Weigt L.A.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF320576; RAQ14805.1; -
KW Hypothetical protein.
SQ SEQUENCE 37 AA; 4308 MW; 551EF5EA0BAA0CAD CRC64;

Query Match 25.6%; Score 51.5; DB 2; Length 37;
Best Local Similarity 38.2%; Pred. No. 54;
Matches 13; Conservative 3; Mismatches 13; Indels 5; Gaps 1;

QY 3 VTSSTFSKLRQLG-----PVTQEFWDNLEKSTEG 31
Db 2 IEKVFFKLARYLAGKTDTPIDDEFVDNLEKAFKG 35

RESULT 2

Q8WTG3 PRELIMINARY; PRT; 39 AA.
AC Q8WTG3;
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Engrailed (fragment).
OS Pepsis sp. MFV-2001.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Vespoidea;
OC Pompilidae; Pepsis.
OX NCBI_TaxID=173816;
RN [1]
RP SEQUENCE FROM N.A.
RA Whiting M.F.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
DR EMBL; AF275780; AALJ5008.1; -
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR InterPro; IPR009057; Homeodomain_like.
DR Pfam; PF00046; Homeobox; 1.
DR PRINTS; PR00024; HOMEOBOX.

על תפוח, וצדדו, הצדד השני, :

```
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0006633; P:fatty acid biosynthesis; IEA.
DR InterPro; IPR000794; Ketoacyl synth.
DR Pfam; PF00109; ketoacyl-synt; 1.
FT NON_TER 41
SQ SEQUENCE 41 AA; 4668 MW; ECA60F116AB6E729 CRC64;

Query Match 21.4%; Score 43; DB 2; Length 41;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 QEFWDNL 25
Db 20 QEFWDNL 26

RESULT 7
Q53103 PRELIMINARY; PRT; 32 AA.
AC Q53103;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE CcOP (Fragment).
OS Rhodobacter sphaeroides (Rhodospseudomonas sphaeroides).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;
OC Rhodobacteraceae; Rhodobacter.
OX NCBI_TaxID=1063;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2.4.1;
RX MEDLINE=96062223; PubMed=7592416;
RA Zeilstra-Ryalls J.H., Kaplan S.;
RT "Aerobic and anaerobic regulation in Rhodobacter sphaeroides 2.4.1:
RT the role of the fnrL gene.";
RL J. Bacteriol. 177:6422-6431 (1995).
DR EMBL; U35445; AAC43501.1; -.
FT NON_TER 32
SQ SEQUENCE 32 AA; 3543 MW; 28D4595324F75D87 CRC64;

Query Match 20.4%; Score 41; DB 2; Length 32;
Best Local Similarity 42.3%; Pred. No. 8.6e+02;
Matches 11; Conservative 2; Mismatches 11; Indels 2; Gaps 1;

QY 2 SVTSTFSKLREQIGPVTOEFWDNLEK 27
Db 2 SVKPTKQKGP-PTGPFWDGIEE 25

RESULT 8
O62886 PRELIMINARY; PRT; 40 AA.
AC O62886;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE MHC class II antigen DLA-DQB1 (Fragment).
GN Name=DLA-DQB1;
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Irish setter;
RA Polvi A., Garden O.A., Batt R.M., Partanen J.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AF043492; AAC05827.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0045012; F:MHC class II receptor activity; IEA.
DR GO; GO:0019884; P:antigen presentation, exogenous antigen; IEA.
DR GO; GO:0019886; P:antigen processing, exogenous antigen via M. . .; IEA.

DR GO; GO:0006955; P:immune response; IEA.
DR InterPro; IPR000353; MHC II beta.
DR Pfam; PF00969; MHC II beta; 1.
DR ProDom; PD000328; MHC_II_beta; 1.
KW Glycoprotein; MHC II; Transmembrane.
FT NON_TER 1
SQ SEQUENCE 40 AA; 4849 MW; 21FA6215B27D5232 CRC64;

Query Match 19.9%; Score 40; DB 2; Length 40;
Best Local Similarity 29.2%; Pred. No. 1.5e+03;
Matches 7; Conservative 7; Mismatches 10; Indels 0; Gaps 0;

QY 13 QLGPVTOEFWDNLEKETEGLRQEM 36
Db 17 ELGRPDAEYVNRQKDEMDRVRAEL 40

RESULT 9
P97599 PRELIMINARY; PRT; 29 AA.
AC P97599;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE DNA primase small subunit (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA Lu J.X.;
RL Submitted (AUG-1996) to the EMBL/GenBank/DBSJ databases.
DR EMBL; U67994; AAB39619.1; -.
FT NON_TER 1
SQ SEQUENCE 29 AA; 3425 MW; 18EF7DD0B8C2CB42 CRC64;

Query Match 19.7%; Score 39.5; DB 2; Length 29;
Best Local Similarity 30.8%; Pred. No. 1.2e+03;
Matches 8; Conservative 8; Mismatches 7; Indels 3; Gaps 1;

QY 17 VTQEFWDNLEKETEGL---RQEMSKD 39
Db 3 VFEQFLENLDKSRKCALLKKSIDLQKD 28

RESULT 10
O45208 PRELIMINARY; PRT; 36 AA.
AC O45208;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Engrailed protein (Fragment).
OS Acanthokara kaputensis.
OC Eukaryota; Metazoa; Onychophora; Acanthokara.
OX NCBI_TaxID=62873;
RN [1]
RP SEQUENCE FROM N.A.
RA Weeden C.J., Kostriken R., Leach D., Whittington P.M.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBSJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Weeden C.J.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBSJ databases.
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
DR EMBL; Y10771; CAA71745.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003709; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
```

```
DR InterPro; IPR009057; Homeodomain_like.
DR Pfam; PF00046; Homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR PROSITE; PS00071; HOMEBOX_2; 1.
DR DNA-binding; Homeobox; Nuclear protein.
FW NON_TER 1
FT NON_TER 36
SQ SEQUENCE 36 AA; 4478 MW; E803EA629E459B67 CRC64;

Query Match 19.7%; Score 39.5; DB 2; Length 36;
Best Local Similarity 39.3%; Pred. No. 1.5e+03;
Matches 11; Conservative 6; Mismatches 10; Indels 1; Gaps 1;

QY 12 EQLGPVTQEFWNLEKETEGLEQNSKD 39
   : : : : : : : : : : : : : : : :
Db 3 EQLRLKKEFQEN-RYLTERKRODLAND 29

RESULT 11
Q9MHZ4 PRELIMINARY; PRT; 27 AA.
AC Q9MHZ4;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Cytochrome b (Fragment).
GN Name=mt:Cyt-b; Synonyms=cytb;
OS Drosophila elegans (Fruit fly).
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=30023;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=12927135;
RA Kastanis P., Eliopoulos E., Goulielmos G.N., Taakas S., Loukas M.;
RT "Macroevolutionary relationships of species of Drosophila melanogaster
group based on mtDNA sequences.";
RL Mol. Phylogenet. Evol. 28:518-528 (2003).
DR EMBL; AF164596; AAF81402.1; -.
DR FlyBase; FBgn0042070; Dele\mt:Cyt-b.
DR GO; GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion.
FT NON_TER 1
SQ SEQUENCE 27 AA; 3300 MW; D772900B152680C8 CRC64;

Query Match 19.4%; Score 39; DB 2; Length 27;
Best Local Similarity 41.7%; Pred. No. 1.2e+03;
Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 14 LGPVTQEFWDNL 25
   : : : : :
Db 14 INPLVSKWWDNL 25

RESULT 12
Q9MHZ6 PRELIMINARY; PRT; 27 AA.
AC Q9MHZ6;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Cytochrome b (fragment).
GN Name=mt:Cyt-b; Synonyms=cytb;
OS Drosophila eugracilis (Fruit fly).
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=29029;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=12927135;
RA Kastanis P., Eliopoulos E., Goulielmos G.N., Taakas S., Loukas M.;
RT "Macroevolutionary relationships of species of Drosophila melanogaster
group based on mtDNA sequences.";
RL Mol. Phylogenet. Evol. 28:518-528 (2003).
DR EMBL; AF164596; AAF81402.1; -.
DR FlyBase; FBgn0042070; Dele\mt:Cyt-b.
DR GO; GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion.
FT NON_TER 1
SQ SEQUENCE 27 AA; 3300 MW; D772900B152680C8 CRC64;

Query Match 19.4%; Score 39; DB 2; Length 27;
Best Local Similarity 41.7%; Pred. No. 1.2e+03;
Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 14 LGPVTQEFWDNL 25
   : : : : :
Db 14 INPLVSKWWDNL 25

RESULT 13
Q9MHZ8 PRELIMINARY; PRT; 27 AA.
AC Q9MHZ8;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Cytochrome b (Fragment).
GN Name=mt:Cyt-b; Synonyms=cytb;
OS Drosophila ficusphila (Fruit fly).
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=30025;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=12927135;
RA Kastanis P., Eliopoulos E., Goulielmos G.N., Taakas S., Loukas M.;
RT "Macroevolutionary relationships of species of Drosophila melanogaster
group based on mtDNA sequences.";
RL Mol. Phylogenet. Evol. 28:518-528 (2003).
DR EMBL; AF164594; AAF81398.1; -.
DR FlyBase; FBgn0042053; Dfic\mt:Cyt-b.
DR GO; GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion.
FT NON_TER 1
SQ SEQUENCE 27 AA; 3300 MW; D772900B035C3C18 CRC64;

Query Match 19.4%; Score 39; DB 2; Length 27;
Best Local Similarity 41.7%; Pred. No. 1.2e+03;
Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 14 LGPVTQEFWDNL 25
   : : : : :
Db 14 INPLVSKWWDNL 25

RESULT 14
Q9MI20 PRELIMINARY; PRT; 27 AA.
AC Q9MI20;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Cytochrome b (Fragment).
GN Name=mt:Cyt-b; Synonyms=cytb;
OS Drosophila kikkawai (Fruit fly).
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
```

```
OX NCBI_TaxID=30033;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=12927135;
RA Kastanis P., Eliopoulos E., Goulielmos G.N., Tsakas S., Loukas M.;
RT "Macroevolutionary relationships of species of Drosophila melanogaster
RT group based on mtDNA sequences.";
RL Mol. Phylogenet. Evol. 28:518-528(2003).
DR EMBL; AF164583; AAF81376.1; -.
DR FlyBase; FBgn0042046; Dkik\mt:Cyt-b.
DR GO; GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion.
FT NON_TER 1 1
SQ SEQUENCE 27 AA; 3313 MW; CD5B900B1530EC18 CRC64;

Query Match 19.4%; Score 39; DB 2; Length 27;
Best Local Similarity 41.7%; Pred. No. 1.2e+03;
Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 14 LGPVTQEFWDNL 25
: | : : |||
Db 14 INPLVKNKWDNL 25

RESULT 15
Q9MI26 PRELIMINARY; PRT; 27 AA.
AC Q9MI26;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cytochrome b (Fragment).
GN Name=mt:Cyt-b; Synonyms=cytb;
OS Drosophila autaria (Fruit fly).
OC Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=47315;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=12927135;
RA Kastanis P., Eliopoulos E., Goulielmos G.N., Tsakas S., Loukas M.;
RT "Macroevolutionary relationships of species of Drosophila melanogaster
RT group based on mtDNA sequences.";
RL Mol. Phylogenet. Evol. 28:518-528(2003).
DR EMBL; AF164580; AAF81370.1; -.
DR FlyBase; FBgn0042077; Daur\mt:Cyt-b.
DR GO; GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion.
FT NON_TER 1 1
SQ SEQUENCE 27 AA; 3313 MW; CD5B900B1530EC18 CRC64;

Query Match 19.4%; Score 39; DB 2; Length 27;
Best Local Similarity 41.7%; Pred. No. 1.2e+03;
Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 14 LGPVTQEFWDNL 25
: | : : |||
Db 14 INPLVKNKWDNL 25

Search completed: December 21, 2004, 12:18:24
Job time : 192 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2004, 12:05:32 ; Search time 151 Seconds
(without alignments)

92.652 Million cell updates/sec

Title: US-09-803-918A-2_COPY_75_113

Perfect score: 201

Sequence: 1 DSVTSTFSKRLQGLGVTFQDFWDLNLEKTEGLRQEMSKD 39

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 856990

Minimum DB seq length: 0

Maximum DB seq length: 41

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_23Sep04:*

1: Geneseqp1980s:*

2: Geneseqp1990s:*

3: Geneseqp2000s:*

4: Geneseqp2001s:*

5: Geneseqp2002s:*

6: Geneseqp2003as:*

7: Geneseqp2003bs:*

8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	89	44.3	16	6	ABP57175 Breast ca
2	80	39.8	16	4	AAG62607 Apolipop
3	60.5	30.1	37	4	AAE02141 Mouse alp
4	60.5	30.1	37	5	AAU99633 Mouse sec
5	56	27.9	22	2	AAV18752 Lecithin:
6	56	27.9	22	2	AAV19006 Lecithin:
7	56	27.9	22	2	AAV19260 Lecithin:
8	56	27.9	22	2	AAV18489 Lecithin:
9	56	27.9	22	8	ADG20969 Apolipop
10	56	27.9	22	8	ADJ32911 Apo lipop
11	55	27.4	31	8	ADP80870 Mouse apo
12	52	25.9	18	7	ADC29662 Antioxi
13	52	25.9	18	7	ADC29658 Antioxi
14	52	25.9	18	7	ADC29661 Antioxi
15	52	25.9	18	7	ADC29657 Antioxi
16	52	25.9	18	7	ADC29659 Antioxi
17	52	25.9	18	7	ADC29660 Antioxi
18	52	25.9	30	7	ADC29669 Antioxi
19	52	25.9	30	7	ADC29665 Antioxi
20	52	25.9	30	7	ADC29668 Antioxi
21	52	25.9	30	7	ADC29670 Antioxi
22	52	25.9	30	7	ADC29664 Antioxi
23	52	25.9	30	7	ADC29667 Antioxi
24	52	25.9	30	7	ADC29666 Antioxi
25	51	25.4	22	2	AAR48545 Sequence

ALIGNMENTS

RESULT 1

ABP57175

ID ABP57175 standard; peptide; 16 AA.

XX AC ABP57175;

XX DT 16-APR-2003 (first entry)

XX DE Breast cancer associated tryptic digest peptide SEQ ID NO:37.

XX KW Breast cancer associated feature; BF; BPI; breast cancer; diagnosis;
breast cancer associated protein isoform; cytostatic; gene therapy.

XX OS Homo sapiens.

XX OS Synthetic.

XX PN WO200288750-A2.

XX PD 07-NOV-2002.

XX PF 02-MAY-2002; 2002WO-GB002022.

XX PR 02-MAY-2001; 2001GB-00010790.

XX PR 27-JUL-2001; 2001GB-00018385.

XX PR 14-AUG-2001; 2001GB-00019791.

XX PR 16-AUG-2001; 2001GB-00020045.

XX PR 22-NOV-2001; 2001GB-00028062.

XX PA (OXFO-) OXFORD GLYSCSCIENCES UK LTD.

XX PI Herath HMAG;

XX DR WPI; 2003-175048/17.

XX PT Screening, diagnosing or determining the stage or severity of breast
cancer, comprises analyzing and quantitatively detecting Breast Cancer-
Associated Features or Breast Cancer-Associated Protein Isoforms in a
biological sample.

XX PS Disclosure; Page 18; 88pp; English.

XX CC The present invention describes a method for screening, diagnosing or
determining the stage or severity of breast cancer, identifying a subject
at risk of developing breast cancer, or monitoring the effect of therapy
administered to a subject with breast cancer, by generating a two-
dimensional array of features comprising breast cancer-associated
features (BPFs), or quantitatively detecting breast cancer-associated
protein isoforms (BPIs). Also described: (1) an antibody capable of

Add94064 Human apo
Aay18754 Lecithin:
Aay19008 Lecithin:
Aay19262 Lecithin:
Aay18491 Lecithin:
Adg20971 Apolipop
Adj32913 Apo lipop
Aay18799 Lecithin:
Aay19053 Lecithin:
Aay19307 Lecithin:
Aay18536 Lecithin:
Adg21016 Apolipop
Adj32958 Apo lipop
Aay18757 Lecithin:
Aay18801 Lecithin:
Aay18756 Lecithin:
Aay19011 Lecithin:
Aay19055 Lecithin:
Aay19010 Lecithin:
Aay19309 Lecithin:

CC immunospecifically binding to one of the BPIs; (2) a pharmaceutical
 CC compositions comprising: (a) a BPI, or a nucleic acid encoding a BPI, and
 CC a carrier; or (b) the antibody of (1), or a fragment or derivative of the
 CC antibody; and a carrier; (3) screening for agents that interact with one
 CC or more BPIs, BPI fragments, polypeptides related to BPIs, or BPI-fusion
 CC proteins; (4) screening for or identifying agents that modulate the
 CC expression or activity of one or more BPIs, a BPI fragment, a BPI-related
 CC polypeptide, or BPI-fusion proteins; and (5) treating or preventing
 CC breast cancer. BPIs have cytostatic activity and can be used in gene
 CC therapy. Methods and kits comprising antibodies or the BPIs from the
 CC present invention can be used for screening, diagnosing or determining
 CC the stage or severity of breast cancer, identifying a subject at risk of
 CC developing breast cancer, or monitoring the effect of therapy
 CC administered to a subject with breast cancer. The antibodies, BPIs,
 CC nucleic acids encoding the BPIs, or an agent that modulates the activity
 CC of one or more BPIs are useful for treating or preventing breast cancer.
 CC ABP57104 to ABP57250 represent breast cancer associated tryptic digest
 CC peptides, which are used in the exemplification of the present invention
 XX
 SQ Sequence 16 AA;

Query Match 44.3%; Score 89; DB 6; Length 16;
 Best Local Similarity 100.0%; Pred. NO. 0.00014;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 EQLGPTQEFWDNLEK 27
 |||||
 Db 1 EQLGPTQEFWDNLEK 16

RESULT 2
 AAG62607
 ID AAG62607 standard; peptide; 16 AA.

XX
 AC AAG62607;

DT 06-SEP-2001 (first entry)

XX Apolipoprotein fragment #1.

XX Apolipoprotein; ApoA-1; atherosclerosis; coronary disease;
 KW cardiovascular disease; ischaemic heart disease; dyslipidaemia.

XX Unidentified.

XX Key Location/Qualifiers

FT Modified-site 14

FT /label= OTHER
 FT /note= "optionally oxidised"

XX WO200138395-A1.

XX 31-MAY-2001.

XX 27-NOV-2000; 2000WO-AU0001463.

XX 26-NOV-1999; 99AU-00004293.

XX (HEAR-) HEART RES INST LTD.

XX Stocker R, Wang XL, Wilcken D;

XX WPI; 2001-355909/37.

XX Novel oxidized form of apolipoprotein useful in diagnosis and treatment
 PT of diseases associated with oxidative stress such as cardiovascular
 PT diseases, in particular, atherosclerosis.

XX Example 1; Page 28; 55pp; English.

XX The present invention relates to oxidised apolipoprotein A-I (ApoA-I)
 CC where at least Met residue 86 is oxidised to Met(O). This can be used in
 CC the prevention, diagnosis and treatment of lipid associated disorders,

CC including coronary vascular disease, ischaemic heart disease,
 CC atherosclerosis and dyslipidaemias. The present sequence is a fragment of
 CC the ApoA-I protein isolated in the exemplification of the invention
 XX
 SQ Sequence 16 AA;

Query Match 39.8%; Score 80; DB 4; Length 16;
 Best Local Similarity 100.0%; Pred. No. 0.002;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 DNLEKETEGLRQEMSK 38
 |||||
 Db 1 DNLEKETEGLRQEMSK 16

RESULT 3
 AAE02141
 ID AAE02141 standard; peptide; 37 AA.

XX AAE02141;

XX 31-JUL-2001 (first entry)

XX Mouse alpha helical protein-34 (Zalpa34) antigenic epitope #12.

XX Mouse; alpha helical protein-34; Zalpa34; antiinfertility;
 KW antigenic epitope; spermatogenesis; educational kit; therapy;
 KW tumour associated antigen.

XX Mus musculus.

XX WO200132884-A2.

XX 10-MAY-2001.

XX 24-OCT-2000; 2000WO-US029277.

XX 29-OCT-1999; 99US-00430153.

XX (ZYMO) ZYMOGENETICS INC.

XX Conklin DC, Taft DW;

XX WPI; 2001-316446/33.

XX Claim 9; Page 66; 77pp; English.

XX The present sequence is mouse alpha helical protein-34 (Zalpa34)
 CC antigenic epitope. Zalpa34 polynucleotides, polypeptides and antibodies
 CC are useful as educational tools, as a laboratory practicum kits for
 CC courses related to genetics and molecular biology, protein chemistry and
 CC antibody production and analysis. Zalpa34 is useful for promoting
 CC spermatogenesis, to identify inhibitors of its activity and to prepare
 CC antibodies that specifically binds to Zalpa34 groups, peptides or
 CC polypeptides. The antibodies are useful for inhibiting spermatogenesis
 CC and to isolate DNA sequences that encode human Zalpa34 genes from cDNA
 CC libraries. Zalpa34 polypeptides are useful as an aid to teach
 CC preparation of antibodies, identifying proteins by western blotting,
 CC protein purification and determining the weight of expressed Zalpa34
 CC polypeptides as a ratio to total protein expressed. Fusion proteins
 CC comprising Zalpa34 and a Fc group are useful as an in vitro assay tool
 CC and antibody-Zalpa34 fusion comprising antibody variable domains are
 CC useful as therapeutic proteins, in which the antibody group binds with
 CC the target antigen, such as a tumour associated antigen

XX Sequence 37 AA;

Query Match 30.1%; Score 60.5; DB 4; Length 37;

Best Local Similarity 54.2%; Pred. No. 1.8;		Matches 13; Conservative 4; Mismatches 6; Indels 1; Gaps 1;	
QY	12 EOLGPVTQEFWDLNLEKETGLRQE 35		
Db	1 BELGE-AQTWDLNQLKELDLREE 23		
RESULT 4			
AAU99633			
ID	AAU99633 standard; peptide; 37 AA.		
XX	AAU99633;		
DT	07-OCT-2002 (first entry)		
XX	Mouse secreted alpha protein 34, Zalpha34, epitope #12.		
DE			
XX	Mouse; secreted alpha protein 34; Zalpha34; epitope; immunogen;		
KW	chromosome 7; spermatogenesis; hormonal; cytokine; teaching aid;		
KW	protein chemistry; enzyme-linked immunoabsorbent assay; ELISA.		
XX			
OS	Mus musculus.		
XX			
FN	US6380361-B1.		
XX			
PD	30-APR-2002.		
XX			
PF	24-OCT-2000; 2000US-00695458.		
XX			
PR	29-OCT-1999; 99US-0162623P.		
XX	(ZYMO) ZYMOGENETICS INC.		
PA			
XX	Conklin DC, Taft DW;		
PI			
XX	WPI; 2002-526165/56.		
DR			
XX	Educational kits and aid for teaching protein chemistry comprising the		
PT	isolated hormonal/cytokine cytokine alpha helical protein-34 (Z alpha 34)		
PT	polypeptide, useful for teaching e.g. ELISA techniques and		
PT	immunopurification.		
XX			
PS	Disclosure; Col 51-52; 35pp; English.		
XX			
CC	The invention relates to an educational kit or aid for teaching protein		
CC	chemistry comprising a container having the isolated, novel hormonal/		
CC	cytokine alpha helical protein-34 (secreted protein alpha, Zalpha34)		
CC	polypeptide, comprising defined amino acid sequences (AAU99639-AAU99654)		
CC	given in the specification and/or the isolated polypeptides comprising		
CC	(AAU99639-AAU99654). The educational kit or teaching aid may be used to		
CC	teach protein chemistry. A novel full-length Zalpha34 protein can be used		
CC	so the student can have practical learning experiences in protein		
CC	purification procedures, protein re-folding when the protein is produced		
CC	in prokaryotic host cells, and how to make both monoclonal and polyclonal		
CC	antibodies. Antibodies that bind to mature Zalpha34 proteins can also be		
CC	used to learn how to make affinity purification columns, to do enzyme-		
CC	linked immunoabsorbent assays (ELISA). The mouse gene encoding Zalpha34		
CC	is located on chromosome 7. Zalpha34 is useful for promoting		
CC	spermatogenesis. The present sequence is an immunogenic epitope from the		
CC	mouse Zalpha34 protein		
XX			
SQ	Sequence 37 AA;		
Query Match 30.1%; Score 60.5; DB 5; Length 37;			
Best Local Similarity 54.2%; Pred. No. 1.8;			
Matches 13; Conservative 4; Mismatches 6; Indels 1; Gaps 1;			
QY	12 EOLGPVTQEFWDLNLEKETGLRQE 35		
Db	1 BELGE-AQTWDLNQLKELDLREE 23		
RESULT 5			
AAU18752			
ID	AAU18752 standard; peptide; 22 AA.		
XX	AAU18752;		
DT	09-JUL-1999 (first entry)		
XX	Lecithin:cholesterol acyltransferase activation exhibiting peptide #73.		
DE			
XX	Apolipoprotein A-I; agonist; dyslipidemic disorder; dyslipidemia; human;		
KW	lecithin:cholesterol acyltransferase; LCAT; hypercholesterolaemia;		
KW	cardiovascular disease; atherosclerosis; restenosis; HDL; apoA-I;		
KW	high density lipoprotein; hypertriglyceridemia; metabolic syndrome;		
KW	septic shock.		
XX			
OS	Synthetic.		
XX	Homo sapiens.		
XX			
PN	WO9916408-A2.		
XX			
PD	08-APR-1999.		
XX			
PF	28-SEP-1998; 98WO-US020328.		
XX			
PR	29-SEP-1997; 97US-00940093.		
XX	(DASS/) DASSEUX J.		
PA	(SEKU/) SEKUL R.		
PA	(BUT/) BUTTNER K.		
PA	(CORN/) CORNUT I.		
PA	(METZ/) METZ G.		
XX			
PI	Dasseux J, Sekul R, Buttner K, Cornut I, Metz G;		
XX			
DR	WPI; 1999-277031/23.		
XX			
PT	Peptide agonists of apolipoprotein A-I.		
XX			
PS	Example; Page 107; 152pp; English.		
XX			
CC	The present invention describes an agonist (A) of apolipoprotein A-I		
CC	(apoA-I) which is a 14-22 residue peptide, or analog, that forms an		
CC	amphipathic alpha-helix in presence of lipids. (A), and their lipid		
CC	complexes, are used to treat or prevent diseases associated with		
CC	dyslipidemia, specifically hypercholesterolaemia, cardiovascular disease,		
CC	atherosclerosis, restenosis, HDL (high density lipoprotein) or apoA-I		
CC	deficiency; hypertriglyceridemia and metabolic syndrome, also for		
CC	treating septic shock. When labeled, (A) can also be used diagnostically		
CC	to measure serum levels of HDL, in particular the HDL subpopulation that		
CC	is involved in retrograde cholesterol transport, also to image HDL at		
CC	e.g. atherosclerotic streaks, and to raise antibodies. AAY18680 to		
CC	AAU18933 represent lecithin:cholesterol acyltransferase (LCAT) activity		
CC	exhibiting core peptides, which are apoA-I agonists		
XX			
SQ	Sequence 22 AA;		
Query Match 27.9%; Score 56; DB 2; Length 22;			
Best Local Similarity 42.9%; Pred. No. 3.9;			
Matches 9; Conservative 5; Mismatches 7; Indels 0; Gaps 0;			
QY	16 PVTQEFWDLNLEKETGLRQEM 36		
Db	1 PVLDEFWEKLNEXLEALKQKL 21		
RESULT 6			
AAU19006			
ID	AAU19006 standard; peptide; 22 AA.		
XX	AAU19006;		
DT	09-JUL-1999 (first entry)		

XX DE Lecithin:cholesterol acyltransferase activation exhibiting peptide #73.
XX KW Apolipoprotein A-I; agonist; dyslipidemic disorder; dyslipidemia; human;
XX KW lecithin:cholesterol acyltransferase; LCAT; hypercholesterolaemia;
KW cardiovascular disease; atherosclerosis; restenosis; HDL; apoA-I;
KW high density lipoprotein; hypertriglyceridemia; metabolic syndrome;
septic shock.
XX OS Synthetic.
OS Homo sapiens.
XX WO9916458-A1.
XX PD 08-APR-1999.
XX PF 28-SEP-1998; 98WO-US020326.
XX PR 29-SEP-1997; 97US-00940096.
XX PA (DASS/) DASSEUX J.
PA (SEKU/) SEKUL R.
PA (BUTT/) BUTTNER K.
PA (CORN/) CORNUT I.
PA (METZ/) METZ G.
XX PI Dasseux J, Sekul R, Buttner K, Cornut I, Metz G;
XX WPI; 1999-277034/23.
XX PT Peptide agonists of apolipoprotein A-I.
XX PS Example; Page 110; 254pp; English.
XX CC The present invention describes an agonist (A) of apolipoprotein A-I
CC (apoA-I) which is a 15-29 residue peptide, or analog, that forms an
CC amphipathic alpha-helix in presence of lipids. (A), and their lipid
CC complexes, are used to treat or prevent diseases associated with
CC dyslipidemia, specifically hypercholesterolaemia, cardiovascular disease,
CC atherosclerosis, restenosis, HDL (high density lipoprotein) or apoA-I
CC deficiency; hypertriglyceridemia and metabolic syndrome, also for
CC treating septic shock. When labeled, (A) can also be used diagnostically
CC to measure serum levels of HDL, in particular the HDL subpopulation that
CC is involved in retrograde cholesterol transport, also to image HDL at
CC e.g. atherosclerotic streaks, and to raise antibodies. AAY18334 to
CC AAY19187 represent lecithin:cholesterol acyltransferase (LCAT) activity
CC exhibiting core peptides, which are apoA-I agonists
XX SQ Sequence 22 AA;
Query Match 27.9%; Score 56; DB 2; Length 22;
Best Local Similarity 42.9%; Pred. No. 3.9;
Matches 9; Conservative 5; Mismatches 7; Indels 0; Gaps 0;
QY 16 PVTQFWDNLEKETEGLRQEM 36
Db 1 PVLDFWEKLNEXLEALKQKL 21
RESULT 7
AAY19260
ID AAY19260 standard; peptide; 22 AA.
XX AAY19260;
XX 14-JUL-1999 (first entry)
XX Lecithin:cholesterol acyltransferase activation exhibiting peptide #73.
XX Apolipoprotein A-I; agonist; dyslipidemic disorder; dyslipidemia; human;
KW lecithin:cholesterol acyltransferase; LCAT; hypercholesterolaemia;
KW cardiovascular disease; atherosclerosis; restenosis; HDL; apoA-I;
KW high density lipoprotein; hypertriglyceridemia; metabolic syndrome;

KW septic shock.
XX Synthetic.
OS Homo sapiens.
XX WO9916459-A1.
XX PD 08-APR-1999.
XX PF 28-SEP-1998; 98WO-US020327.
XX PR 29-SEP-1997; 97US-00940095.
XX PA (DASS/) DASSEUX J.
PA (SEKU/) SEKUL R.
PA (BUTT/) BUTTNER K.
PA (CORN/) CORNUT I.
PA (METZ/) METZ G.
PA (DUFO/) DUFOURCQ J.
XX PI Dasseux J, Sekul R, Buttner K, Cornut I, Metz G, Dufourcq J;
XX WPI; 1999-277035/23.
XX PT Peptide agonists of apolipoprotein A-I.
XX PS Example; Page 119; 280pp; English.
XX CC The present invention describes an agonist (A) of apolipoprotein A-I
CC (apoA-I) which is a 15-29 residue peptide, or analog, that forms an
CC amphipathic alpha-helix in presence of lipids. (A), and their lipid
CC complexes, are used to treat or prevent diseases associated with
CC dyslipidemia, specifically hypercholesterolaemia, cardiovascular disease,
CC atherosclerosis, restenosis, HDL (high density lipoprotein) or apoA-I
CC deficiency; hypertriglyceridemia and metabolic syndrome, also for
CC treating septic shock. When labeled, (A) can also be used diagnostically
CC to measure serum levels of HDL, in particular the HDL subpopulation that
CC is involved in retrograde cholesterol transport, also to image HDL at
CC e.g. atherosclerotic streaks, and to raise antibodies. AAY19188 to
CC AAY19441 represent lecithin:cholesterol acyltransferase (LCAT) activity
CC exhibiting core peptides, which are apoA-I agonists
XX SQ Sequence 22 AA;
Query Match 27.9%; Score 56; DB 2; Length 22;
Best Local Similarity 42.9%; Pred. No. 3.9;
Matches 9; Conservative 5; Mismatches 7; Indels 0; Gaps 0;
QY 16 PVTQFWDNLEKETEGLRQEM 36
Db 1 PVLDFWEKLNEXLEALKQKL 21
RESULT 8
AAY18489
ID AAY18489 standard; peptide; 22 AA.
XX AAY18489;
XX 09-JUL-1999 (first entry)
XX Lecithin:cholesterol acyltransferase activation exhibiting peptide #73.
XX Gene therapy; apolipoprotein A-I; agonist; dyslipidemic disorder; ApoA-I;
KW cardiovascular disease; atherosclerosis; restenosis; LCAT;
KW hyperlipidemia; septic shock; lecithin:cholesterol acyltransferase.
XX Synthetic.
OS Homo sapiens.
XX WO9916409-A2.
XX PD 08-APR-1999.

PI Dasseux J, Sekul R, Buttner K, Cornut I, Metz G;
 XX WPI; 2004-059110/06.
 XX
 PT New Apo lipoprotein A-I agonist useful for treating disorders associated
 PT with dyslipoproteinemia e.g. hypercholesterolemia or septic shock,
 PT comprises 15-29 residue peptide or analog which forms amphipathic alpha-
 PT helix in presence of lipids.
 XX
 XX Example 8; SEQ ID NO 73; 135pp; English.
 XX
 CC The invention comprises Apo lipoprotein A-I (ApoA-I) agonist peptides
 CC which form an amphipathic alpha-helix in the presence of lipids. The ApoA
 CC -I peptides of the invention are useful for treating a disorder
 CC associated with dyslipidaemia, such as: hypercholesterolaemia,
 CC cardiovascular disease, atherosclerosis, restenosis, high density
 CC lipoprotein (HDL) deficiency, ApoA-I deficiency, hypertriglyceridaemia,
 CC metabolic syndrome, and septic shock. The ApoA-I peptides are also useful
 CC for treating coronary heart disease and endotoxaemia. The present amino
 CC acid sequence represents an ApoA-I agonist peptide of the invention.
 XX
 XX Sequence 22 AA;
 SQ
 Query Match 27.9%; Score 56; DB 8; Length 22;
 Best Local Similarity 42.9%; Pred. No. 3.9;
 Matches 9; Conservative 5; Mismatches 7; Indels 0; Gaps 0;
 QY 16 PVTQFWDNLEKETGLRQEM 36
 Db |||||:|:|:|:|:
 1 PVLDFEWEKLNEXLEALKQKL 21
 RESULT 11
 ADP80870
 ID ADP80870 standard; peptide; 31 AA.
 AC
 XX ADP80870;
 XX
 DT 09-SEP-2004 (first entry)
 XX
 DE Mouse apolipoprotein E amino acid sequence SEQ ID NO:95.
 XX
 KW c-fos; c-fos interacting protein; fos interacting protein chromosome X;
 KW Fip-cx; screening; Fc-fos interacting protein inhibitor; mouse;
 KW apolipoprotein E.
 XX
 OS Mus musculus.
 XX
 PN WO2004053121-A1.
 XX
 PD 24-JUN-2004.
 XX
 PF 19-NOV-2003; 2003WO-JP014749.
 XX
 PR 11-DEC-2002; 2002JP-00360046.
 XX
 PA (UYKE-) UNIV KEIO.
 XX
 XX Miyamoto E, Ishizaka M, Yanagawa H;
 PI WPI; 2004-517250/49.
 DR N-PSDB; ADP80934.
 XX
 XX New proteins that interact with fos, e.g., fos interacting protein
 PT chromosome X (Fip-cx).
 PT
 XX Claim 89; SEQ ID NO 95; 192pp; Japanese.
 XX
 CC The present invention describes a protein (I) that interacts with c-fos
 CC (e.g., fos interacting protein chromosome X (Fip-cx), Fip-cx.1, fos
 CC interacting protein chromosome ex.2, or fos interacting protein
 CC chromosome 4). Also described: (1) a nucleic acid (II) that encodes (I);
 CC (2) an inhibitor (III) that inhibits the interaction of (I) encoded by

CC (II) with c-fos protein; (3) an inhibitor (IV) that inhibits interaction
 CC of a protein with c-fos protein; and (4) detecting (M1) the interaction
 CC of a protein as mentioned in (3) with c-fos protein, by contacting the
 CC protein with c-fos, to form a composite. (I) is useful for detecting its
 CC interaction with c-fos, to form a complex. (M1) is useful for screening
 CC the protein that interacts with c-fos which involves performing the
 CC detection process and selecting the protein that interacts with c-fos.
 CC (M1) is useful for screening a protein that interacts with c-fos which
 CC involves performing (M1) and selecting the detected protein. (I) is
 CC useful for screening inhibitors that interact with c-fos. The present
 CC sequence represents a mouse apolipoprotein E amino acid sequence, which
 CC can interact with c-fos in the exemplification of the present invention.
 XX
 XX Sequence 31 AA;
 SQ
 Query Match 27.4%; Score 55; DB 8; Length 31;
 Best Local Similarity 52.2%; Pred. No. 7.8;
 Matches 12; Conservative 3; Mismatches 8; Indels 0; Gaps 0;
 QY 9 KLRQLGPTVQTFWDNLEKETEG 31
 Db :|||||:|:|:|:|:
 9 ELEQLGPEVAEETRARLGKEQG 31
 RESULT 12
 ADC29662
 ID ADC29662 standard; peptide; 18 AA.
 XX
 AC ADC29662;
 XX
 DT 18-DEC-2003 (first entry)
 XX
 DE Antioxidant peptide #37.
 XX
 KW antioxidant; cardiovascular disease; ischaemia; bone disease;
 KW inflammatory related disease; human.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN US2003087819-A1.
 XX
 PD 08-MAY-2003.
 XX
 PF 08-MAY-2002; 2002US-00142238.
 XX
 PR 09-MAY-2001; 2001US-0289944P.
 XX
 PA (BIEL/) BIELICKI J K.
 XX
 XX Bielicki JK;
 PI WPI; 2003-786946/74.
 XX
 DR New cysteine-containing peptides, having antioxidant properties, useful
 PT for diagnosing and treating cardiovascular disease, ischemia, bone
 PT disease and other inflammatory related diseases.
 XX
 XX Example 8; SEQ ID NO 37; 49pp; English.
 PS
 CC The invention relates to an antioxidant peptide. The methods and
 CC compositions of the present invention are useful for diagnosing and
 CC treating cardiovascular disease, ischaemia, bone disease and other
 CC inflammatory related diseases. The present sequence represents the amino
 CC acid sequence of a peptide with antioxidant activity.
 XX
 XX Sequence 18 AA;
 SQ
 Query Match 25.9%; Score 52; DB 7; Length 18;
 Best Local Similarity 100.0%; Pred. No. 10;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DSVTSTFSKLR 11

CC The invention relates to an antioxidant peptide. The methods and
CC compositions of the present invention are useful for diagnosing and
CC treating cardiovascular disease, ischaemia, bone disease and other
CC inflammatory related diseases. The present sequence represents the amino
CC acid sequence of a peptide with antioxidant activity.

XX

SQ Sequence 18 AA;

Query Match 25.9%; Score 52; DB 7; Length 18;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DSVTSTFSKLR 11

Db ||||| |||||

8 DSVTSTFSKLR 18

Search completed: December 21, 2004, 12:21:51
Job time : 155 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2004, 11:55:11 ; Search time 142 Seconds
(without alignments)
98.279 Million cell updates/sec

Title: US-09-803-918A-2_COPY_75_113

Perfect score: 201

Sequence: 1 DSVTFTFKLRQLGPTQSFWDNLEKETGLRQEMSKD 39

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1589859 seqs, 357834939 residues

Total number of hits satisfying chosen parameters: 382105

Minimum DB seq length: 0

Maximum DB seq length: 41

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	118	58.7	22	15	US-10-465-789A-50
2	89	44.3	16	17	US-10-700-340-37
3	72	35.8	22	15	US-10-465-789A-49
4	56	27.9	22	10	US-09-865-989-73
5	56	27.9	22	11	US-09-865-989-73
6	56	27.9	22	14	US-10-099-574A-73
7	56	27.9	22	14	US-10-147-849-73
8	56	27.9	22	14	US-10-099-836B-73
9	56	27.9	22	14	US-10-283-599-73
10	56	27.9	22	16	US-10-802-080-73
11	56	27.9	22	17	US-10-801-897-73
12	53	26.4	11	15	US-10-465-789A-59
13	52	25.9	18	14	US-10-142-238A-32

14	52	25.9	18	14	US-10-142-238A-33	Sequence 33, Appl
15	52	25.9	18	14	US-10-142-238A-34	Sequence 34, Appl
16	52	25.9	18	14	US-10-142-238A-35	Sequence 35, Appl
17	52	25.9	18	14	US-10-142-238A-36	Sequence 36, Appl
18	52	25.9	18	14	US-10-142-238A-37	Sequence 37, Appl
19	52	25.9	30	14	US-10-142-238A-39	Sequence 39, Appl
20	52	25.9	30	14	US-10-142-238A-40	Sequence 40, Appl
21	52	25.9	30	14	US-10-142-238A-41	Sequence 41, Appl
22	52	25.9	30	14	US-10-142-238A-42	Sequence 42, Appl
23	52	25.9	30	14	US-10-142-238A-43	Sequence 43, Appl
24	52	25.9	30	14	US-10-142-238A-44	Sequence 44, Appl
25	52	25.9	30	14	US-10-142-238A-45	Sequence 45, Appl
26	50	24.9	10	10	US-09-991-627-2	Sequence 2, Appl
27	48	23.9	22	10	US-09-865-989-75	Sequence 75, Appl
28	48	23.9	22	11	US-09-865-989-75	Sequence 75, Appl
29	48	23.9	22	14	US-10-099-574A-75	Sequence 75, Appl
30	48	23.9	22	14	US-10-147-849-75	Sequence 75, Appl
31	48	23.9	22	14	US-10-099-836B-75	Sequence 75, Appl
32	48	23.9	22	14	US-10-283-599-75	Sequence 75, Appl
33	48	23.9	22	16	US-10-802-080-75	Sequence 75, Appl
34	48	23.9	22	17	US-10-801-897-75	Sequence 75, Appl
35	47	23.4	22	10	US-09-865-989-120	Sequence 120, App
36	47	23.4	22	11	US-09-865-989-120	Sequence 120, App
37	47	23.4	22	14	US-10-099-574A-120	Sequence 120, App
38	47	23.4	22	14	US-10-147-849-120	Sequence 120, App
39	47	23.4	22	14	US-10-099-836B-120	Sequence 120, App
40	47	23.4	22	14	US-10-283-599-120	Sequence 120, App
41	47	23.4	22	16	US-10-802-080-120	Sequence 120, App
42	47	23.4	22	17	US-10-801-897-120	Sequence 120, App
43	46	22.9	22	10	US-09-865-989-77	Sequence 77, Appl
44	46	22.9	22	10	US-09-865-989-78	Sequence 78, Appl
45	46	22.9	22	10	US-09-865-989-122	Sequence 122, App

ALIGNMENTS

RESULT 1
US-10-465-789A-50
; Sequence 50, Application US/10465789A
; Publication No. US20040053384A1
; GENERAL INFORMATION:
; APPLICANT: Sligar, Stephen G
; APPLICANT: Bayburt, Timothy H
; APPLICANT: Schuler, Mary A
; APPLICANT: Civlian, Natanva R
; APPLICANT: Yelena V. Grinkova
; APPLICANT: Iliia G. Denisov
; TITLE OF INVENTION: Membrane Scaffold Proteins
; FILE REFERENCE: 87-00A
; CURRENT APPLICATION NUMBER: US/10/465,789A
; CURRENT FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: 09/990,087
; PRIOR FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 60/252,233
; PRIOR FILING DATE: 2000-11-20
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 50
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: artificial sequence of Helix 2
US-10-465-789A-50

Query Match 58.7%; Score 118; DB 15; Length 22;
Best Local Similarity 100.0%; Pred. No. 1.4e-08;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 PVTQSFWDNLEKETGLRQEMS 37

Db 1 PVTQSFWDNLEKETGLRQEMS 22

Matches	15;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	1	DSVTSTFSKLREQLG	15						
Db	8	DSVTSTFSKLREQLG	22						
RESULT 4									
US-09-865-989-73									
: Sequence 73, Application US/09865989									
: Publication No. US2003000827A1									
: GENERAL INFORMATION:									
: APPLICANT: Daseux, Jean-Louis									
: Sekul, Renate									
: Buttner, Klaus									
: Cornut, Isabelle									
: Metz, Gunther									
: TITLE OF INVENTION: APOLIPOPROTEIN A-I AGONISTS									
: AND THEIR USE TO TREAT DYSLIPIDEMIC DISORDERS									
: NUMBER OF SEQUENCES: 258									
: CORRESPONDENCE ADDRESS:									
: ADDRESSEE: Pennie & Edmonds LLP									
: STREET: 1155 Avenue of the Americas									
: CITY: New York									
: STATE: NY									
: COUNTRY: USA									
: ZIP: 10036-2811									
: COMPUTER READABLE FORM:									
: MEDIUM TYPE: Diskette									
: COMPUTER: IBM Compatible									
: OPERATING SYSTEM: DOS									
: SOFTWARE: FastSeq Version 2.0									
: CURRENT APPLICATION DATA:									
: APPLICATION NUMBER: US/09/865,989									
: FILING DATE: 25-May-2001									
: CLASSIFICATION: <Unknown>									
: PRIOR APPLICATION DATA:									
: APPLICATION NUMBER: 09/465,719									
: FILING DATE: 17-DEC-1999									
: ATTORNEY/AGENT INFORMATION:									
: NAME: Coruzzi, Laura A									
: REGISTRATION NUMBER: 30,742									
: REFERENCE/DOCKET NUMBER: 009196-0006-999									
: TELECOMMUNICATION INFORMATION:									
: TELEPHONE: 650-493-4935									
: TELEFAX: 650-493-5556									
: TELEX: 66141 PENNIE									
: INFORMATION FOR SEQ ID NO: 73:									
: SEQUENCE CHARACTERISTICS:									
: LENGTH: 22 amino acids									
: TYPE: amino acid									
: STRANDEDNESS: single									
: TOPOLOGY: linear									
: MOLECULE TYPE: No. US2003000827A1e									
: FEATURE:									
: NAME/KEY: Other									
: LOCATION: 13									
: OTHER INFORMATION: Xaa = Aib									
: SEQUENCE DESCRIPTION: SEQ ID NO: 73:									
US-09-865-989-73									
Query Match 27.9%; Score 56; DB 10; Length 22;									
Best Local Similarity 42.9%; Pred. No. 2.3;									
Matches 9; Conservative 5; Mismatches 7; Indels 0; Gaps 0;									

```
Publication No. US20040029807A9
GENERAL INFORMATION:
APPLICANT: Dasseux, Jean-Louis
            Sekul, Renate
            Buttner, Klaus
            Cornut, Isabelle
            Metz, Gunther
TITLE OF INVENTION: APOLIPOPROTEIN A-I AGONISTS
AND THEIR USE TO TREAT DYSLIPIDEMIC DISORDERS
NUMBER OF SEQUENCES: 258
CORRESPONDENCE ADDRESS:
ADDRESSER: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036-2811
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA: US/09/865,989
FILING DATE: 25-May-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/465,719
FILING DATE: 17-DEC-1999
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 009196-0006-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-493-4935
TELEFAX: 650-493-5556
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 73:
SEQUENCE CHARACTERISTICS:
LENGTH: 22 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. US20040029807A9e
FEATURE:
NAME/KEY: Other
LOCATION: 13
OTHER INFORMATION: Xaa = Aib
SEQUENCE DESCRIPTION: SEQ ID NO: 73:
US-09-865-989-73
Query Match 27.9%; Score 56; DB 11; Length 22;
Best Local Similarity 42.9%; Pred. No. 2.3;
Matches 9; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
RESULT 6
US-10-099-574A-73
Sequence 73, Application US/10099574A
Publication No. US20030060604A1
GENERAL INFORMATION:
APPLICANT: Dasseux, Jean-Louis
APPLICANT: Sekul, Renate
APPLICANT: Buttner, Klaus
APPLICANT: Cornut, Isabelle
APPLICANT: Metz, Gunther
TITLE OF INVENTION: APOLIPOPROTEIN A-I AGONISTS
TITLE OF INVENTION: AND THEIR USE TO TREAT DYSLIPIDEMIC DISORDERS
NUMBER OF SEQUENCES: 254
CORRESPONDENCE ADDRESS:
ADDRESSER: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036-2811
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
```

```
Publication No. US20040029807A9
GENERAL INFORMATION:
APPLICANT: Dasseux, Jean-Louis
            Sekul, Renate
            Buttner, Klaus
            Cornut, Isabelle
            Metz, Gunther
TITLE OF INVENTION: APOLIPOPROTEIN A-I AGONISTS
AND THEIR USE TO TREAT DYSLIPIDEMIC DISORDERS
NUMBER OF SEQUENCES: 258
CORRESPONDENCE ADDRESS:
ADDRESSER: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036-2811
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA: US/10/099,574A
FILING DATE: 29-SEP-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 009196-0005-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-493-4935
TELEFAX: 650-493-5556
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 73:
SEQUENCE CHARACTERISTICS:
LENGTH: 22 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. US20030060604A1e
FEATURE:
NAME/KEY: Other
LOCATION: 13
OTHER INFORMATION: Xaa = Aib
US-10-099-574A-73
Query Match 27.9%; Score 56; DB 14; Length 22;
Best Local Similarity 42.9%; Pred. No. 2.3;
Matches 9; Conservative 5; Mismatches 7; Indels 0; Gaps 0;
RESULT 7
US-10-147-849-73
Sequence 73, Application US/10147849
Publication No. US20030190319A1
GENERAL INFORMATION:
APPLICANT: Dasseux, Jean-Louis
APPLICANT: Sekul, Renate
APPLICANT: Buttner, Klaus
APPLICANT: Cornut, Isabelle
APPLICANT: Metz, Gunther
TITLE OF INVENTION: APOLIPOPROTEIN A-I AGONISTS
TITLE OF INVENTION: AND THEIR USE TO TREAT DYSLIPIDEMIC DISORDERS
NUMBER OF SEQUENCES: 254
CORRESPONDENCE ADDRESS:
ADDRESSER: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036-2811
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
```

```
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/147,849
; FILING DATE: 29-SEP-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 009196-0005-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-493-4935
; TELEFAX: 650-493-5556
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 73:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. US20030190319A1e
; FEATURE:
; NAME/KEY: Other
; LOCATION: 13
; OTHER INFORMATION: Xaa = Aib
; US-10-147-849-73

Query Match 27.9%; Score 56; DB 14; Length 22;
Best Local Similarity 42.9%; Pred. No. 2.3;
Matches 9; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 16 PVTQEFWDNLEKETEGLRQEM 36
Db 1 PVLDEFWEKLNEXLEALKQKL 21

RESULT 8
US-10-099-836B-73
; Sequence 73, Application US/10099836B
; Publication No. US20030203842A1
; GENERAL INFORMATION:
; APPLICANT: Dasseux, Jean-Louis
; Sekul, Renate
; Buttner, Klaus
; Cornut, Isabelle
; Metz, Gunther
; Dufourcq, Jean
; TITLE OF INVENTION: APOLIPOPROTEIN A-I AGONISTS
; NUMBER OF SEQUENCES: 254
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/099,836B
; FILING DATE: 28-Aug-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: <Unknown>
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A
```

```
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 009196-0004-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-493-4935
; TELEFAX: 650-493-5556
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 73:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. US20030203842A1e
; FEATURE:
; NAME/KEY: Other
; LOCATION: 13
; OTHER INFORMATION: Xaa = Aib
; SEQUENCE DESCRIPTION: SEQ ID NO: 73:
US-10-099-836B-73

Query Match 27.9%; Score 56; DB 14; Length 22;
Best Local Similarity 42.9%; Pred. No. 2.3;
Matches 9; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 16 PVTQEFWDNLEKETEGLRQEM 36
Db 1 PVLDEFWEKLNEXLEALKQKL 21

RESULT 9
US-10-283-599-73
; Sequence 73, Application US/10283599
; Publication No. US20030208059A1
; GENERAL INFORMATION:
; APPLICANT: Dasseux, Jean-Louis
; APPLICANT: Sekul, Renate
; APPLICANT: Buttner, Klaus
; APPLICANT: Cornut, Isabelle
; APPLICANT: Metz, Gunther
; APPLICANT: Dufourcq, Jean
; TITLE OF INVENTION: GENE THERAPY APPROACHES TO
; SUPPLY APOLIPOPROTEIN A-I AGONISTS AND THEIR
; TITLE OF INVENTION: USE TO TREAT DYSLIPIDEMIC DISORDERS.
; NUMBER OF SEQUENCES: 274
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/283,599
; FILING DATE: 29-OCT-2002
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/940,136
; FILING DATE: 29-SEP-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 009196-0007-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-493-4935
; TELEFAX: 650-493-5556
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 73:
; SEQUENCE CHARACTERISTICS:
```

LENGTH: 22 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. US20030208059A1e
FEATURE:
NAME/KEY: Other
LOCATION: 13
OTHER INFORMATION: Xaa = Aib
US-10-283-599-73

Query Match 27.9%; Score 56; DB 14; Length 22;
Best Local Similarity 42.9%; Pred. No. 2.3;
Matches 9; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 16 PVTQEFWDNLEKETEGLRQEM 36
Db 1 PVLDFWEKLNEXLEALKQKL 21

RESULT 10
US-10-802-080-73
Sequence 73, Application US/10802080
Publication No. US20040181034A1
GENERAL INFORMATION:
APPLICANT: Dasseux, Jean-Louis
Sekul, Renate
Buttner, Klaus
Cornut, Isabelle
Metz, Gunther
Dufourcq, Jean

TITLE OF INVENTION: APOLIPOPROTEIN A-I AGONISTS
AND THEIR USE TO TREAT DYSLIPIDEMIC DISORDERS
NUMBER OF SEQUENCES: 258
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036-2811

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/802,080
FILING DATE: 15-Mar-2004
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/453,840
FILING DATE: 01-DEC-1999
APPLICATION NUMBER: 08/940,095
FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 009196-0004-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-493-4935
TELEFAX: 650-493-5556
TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 73:
SEQUENCE CHARACTERISTICS:
LENGTH: 22 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: None
FEATURE:
NAME/KEY: Other
LOCATION: 13

OTHER INFORMATION: Xaa = Aib
SEQUENCE DESCRIPTION: SEQ ID NO: 73:
US-10-802-080-73

Query Match 27.9%; Score 56; DB 16; Length 22;
Best Local Similarity 42.9%; Pred. No. 2.3;
Matches 9; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 16 PVTQEFWDNLEKETEGLRQEM 36
Db 1 PVLDFWEKLNEXLEALKQKL 21

RESULT 11
US-10-801-897-73
Sequence 73, Application US/10801897
Publication No. US20040198662A1
GENERAL INFORMATION:
APPLICANT: Dasseux, Jean-Louis
Sekul, Renate
Buttner, Klaus
Cornut, Isabelle
Metz, Gunther

TITLE OF INVENTION: APOLIPOPROTEIN A-I AGONISTS
AND THEIR USE TO TREAT DYSLIPIDEMIC DISORDERS
NUMBER OF SEQUENCES: 258
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036-2811

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/801,897
FILING DATE: 15-Mar-2004
CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/865,989
FILING DATE: 25-May-2001
APPLICATION NUMBER: 09/465,719
FILING DATE: 17-DEC-1999
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 009196-0006-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-493-4935
TELEFAX: 650-493-5556
TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 73:
SEQUENCE CHARACTERISTICS:
LENGTH: 22 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: None
FEATURE:
NAME/KEY: Other
LOCATION: 13
OTHER INFORMATION: Xaa = Aib
US-10-801-897-73

Query Match 27.9%; Score 56; DB 17; Length 22;
Best Local Similarity 42.9%; Pred. No. 2.3;
Matches 9; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 16 PVTQEFWDNLEKETEGLRQEM 36

```

Query Match      25.9%; Score 52; DB 14; Length 18;
Best Local Similarity 100.0%; Pred. No. 6.2;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  DSVTSTSTFKSLR 11
          | | | | | | | |
Db      8  DSVTSTSTFKSLR 18
          | | | | | | | |

```

```

Query Match      25.9%; Score 52; DB 14; Length 18;
Best Local Similarity 100.0%; Pred.No. 6.2;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 DSVTSTFSKLR 11
        |||||
Db      A DSVTSTFSKLR 18

```

Search completed: December 21, 2004, 12:04:53
Job time : 143 secs

;; COUNTRY: USA
;; ZIP: 10036-2811
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette
;; COMPUTER: IBM Compatible
;; OPERATING SYSTEM: DOS
;; SOFTWARE: FastSEQ Version 2.0
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/940,095
;; FILING DATE: 29-SEP-1997
;; CLASSIFICATION: 514
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER:
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Coruzzi, Laura A
;; REGISTRATION NUMBER: 30,742
;; REFERENCE/DOCKET NUMBER: 009196-0004-999
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 650-493-4935
;; TELEFAX: 650-493-5556
;; TELEX: 66141 PENNIE
;; INFORMATION FOR SEQ ID NO: 73:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 22 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: No. 6004925e
;; FEATURE:
;; NAME/KEY: Other
;; LOCATION: 13
;; OTHER INFORMATION: Xaa = Aib
;;
US-08-940-095-73

Query Match 27.9%; Score 56; DB 3; Length 22;
Best Local Similarity 42.9%; Pred. No. 0.58;
Matches 9; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 16 PVTQEFWDNLEKETGLRQEM 36
|||:::|:|:|:|:
Db 1 PVLDFEWEKLNEXLEALKQKL 21

RESULT 3
US-08-940-093-73
; Sequence 73, Application US/08940093
; Patent No. 6037323
; GENERAL INFORMATION:
; APPLICANT: Dasseux, Jean-Louis
; APPLICANT: Sekul, Renate
; APPLICANT: Buttner, Klaus
; APPLICANT: Cornut, Isabelle
; APPLICANT: Metz, Gunther
; TITLE OF INVENTION: APOLIPOPROTEIN A-I AGONISTS
; NUMBER OF SEQUENCES: 258
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/940,093
; FILING DATE: 29-SEP-1997
; CLASSIFICATION: 435

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER:
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Coruzzi, Laura A
;; REGISTRATION NUMBER: 30,742
;; REFERENCE/DOCKET NUMBER: 009196-0006-999
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 650-493-4935
;; TELEFAX: 650-493-5556
;; TELEX: 66141 PENNIE
;; INFORMATION FOR SEQ ID NO: 73:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 22 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: No. 6037323e
;; FEATURE:
;; NAME/KEY: Other
;; LOCATION: 13
;; OTHER INFORMATION: Xaa = Aib
;;
US-08-940-093-73

Query Match 27.9%; Score 56; DB 3; Length 22;
Best Local Similarity 42.9%; Pred. No. 0.58;
Matches 9; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 16 PVTQEFWDNLEKETGLRQEM 36
|||:::|:|:|:|:
Db 1 PVLDFEWEKLNEXLEALKQKL 21

RESULT 4
US-08-940-096-73
; Sequence 73, Application US/08940096
; Patent No. 6046166
; GENERAL INFORMATION:
; APPLICANT: Dasseux, Jean-Louis
; APPLICANT: Sekul, Renate
; APPLICANT: Buttner, Klaus
; APPLICANT: Cornut, Isabelle
; APPLICANT: Metz, Gunther
; TITLE OF INVENTION: APOLIPOPROTEIN A-I AGONISTS
; NUMBER OF SEQUENCES: 258
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/940,096
; FILING DATE: 29-SEP-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 009196-0005-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-493-4935
; TELEFAX: 650-493-5556
; TELEX: 66141 PENNIE

; INFORMATION FOR SEQ ID NO: 73:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 6046166e
; FEATURE:
; NAME/KEY: Other
; LOCATION: 13
; OTHER INFORMATION: Xaa = Aib
US-08-940-096-73

Query Match 27.9%; Score 56; DB 3; Length 22;
Best Local Similarity 42.9%; Pred. No. 0.58;
Matches 9; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 16 PVTQEFWDNLEKETEGLRQEM 36
DB 1 PVLDFEWEKLNEXLEALKQKL 21

RESULT 5
US-09-465-719-73
; Sequence 73, Application US/09465719
; Patent No. 6265377
; GENERAL INFORMATION:
; APPLICANT: Dasseux, Jean-Louis
; APPLICANT: Sekul, Renate
; APPLICANT: Buttner, Klaus
; APPLICANT: Cornut, Isabelle
; APPLICANT: Metz, Gunther
; TITLE OF INVENTION: APOLIPOPROTEIN A-I AGONISTS
; NUMBER OF SEQUENCES: 258
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/465,719
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/940,093
; FILING DATE: 29-SEP-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 009196-0006-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-493-4935
; TELEFAX: 650-493-5556
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 73:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 6265377e
; FEATURE:
; NAME/KEY: Other
; LOCATION: 13
; OTHER INFORMATION: Xaa = Aib

US-09-465-719-73

Query Match 27.9%; Score 56; DB 3; Length 22;
Best Local Similarity 42.9%; Pred. No. 0.58;
Matches 9; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 16 PVTQEFWDNLEKETEGLRQEM 36
DB 1 PVLDFEWEKLNEXLEALKQKL 21

RESULT 6
US-09-453-605-73
; Sequence 73, Application US/09453605
; Patent No. 6329341
; GENERAL INFORMATION:
; APPLICANT: Dasseux, Jean-Louis
; APPLICANT: Sekul, Renate
; APPLICANT: Buttner, Klaus
; APPLICANT: Cornut, Isabelle
; APPLICANT: Metz, Gunther
; APPLICANT: Dufourcq, Jean
; TITLE OF INVENTION: APOLIPOPROTEIN A-I AGONISTS
; NUMBER OF SEQUENCES: 258
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/453,605
; FILING DATE: 26-NO. 6329341-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/940,095
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 009196-0004-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-493-4935
; TELEFAX: 650-493-5556
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 73:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 6329341e
; FEATURE:
; NAME/KEY: Other
; LOCATION: 13
; OTHER INFORMATION: Xaa = Aib
; SEQUENCE DESCRIPTION: SEQ ID NO: 73:
US-09-453-605-73

Query Match 27.9%; Score 56; DB 3; Length 22;
Best Local Similarity 42.9%; Pred. No. 0.58;
Matches 9; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 16 PVTQEFWDNLEKETEGLRQEM 36
DB 1 PVLDFEWEKLNEXLEALKQKL 21

;
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/453,841
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/940,095
; FILING DATE: 29-SEP-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 009196-0004-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-493-5556
; TELEFAX: 650-493-5556
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 73:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 6573239e
; FEATURE:
; NAME/KEY: Other
; LOCATION: 13
; OTHER INFORMATION: Xaa = Aib
; US-09-453-841-73

Query Match 27.9%; Score 56; DB 4; Length 22;
Best Local Similarity 42.9%; Pred. No. 0.58;
Matches 9; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 16 PVTQEFWDNLEKETEGLRQEM 36
DB 1 PVLDFEWEKLNEXLEALKQKL 21

RESULT 10
US-09-453-833-73
; Sequence 73, Application US/09453833
; Patent No. 6602854
; GENERAL INFORMATION:
; APPLICANT: Dasseux, Jean-Louis
; APPLICANT: Sekul, Renate
; APPLICANT: Buttner, Klaus
; APPLICANT: Cornut, Isabelle
; APPLICANT: Metz, Gunther
; APPLICANT: Dufourcq, Jean
; TITLE OF INVENTION: APOLIPOPROTEIN A-I AGONISTS
; TITLE OF INVENTION: AND THEIR USE TO TREAT DYSLIPIDEMIC DISORDERS
; NUMBER OF SEQUENCES: 258
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA:

;
; APPLICATION NUMBER: US/09/453,833
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/940,095
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 009196-0004-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-493-4935
; TELEFAX: 650-493-5556
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 73:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 6602854e
; FEATURE:
; NAME/KEY: Other
; LOCATION: 13
; OTHER INFORMATION: Xaa = Aib
; US-09-453-833-73

Query Match 27.9%; Score 56; DB 4; Length 22;
Best Local Similarity 42.9%; Pred. No. 0.58;
Matches 9; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 16 PVTQEFWDNLEKETEGLRQEM 36
DB 1 PVLDFEWEKLNEXLEALKQKL 21

RESULT 11
US-09-453-826-73
; Sequence 73, Application US/09453826
; Patent No. 6630450
; GENERAL INFORMATION:
; APPLICANT: Dasseux, Jean-Louis
; APPLICANT: Sekul, Renate
; APPLICANT: Buttner, Klaus
; APPLICANT: Cornut, Isabelle
; APPLICANT: Metz, Gunther
; APPLICANT: Dufourcq, Jean
; TITLE OF INVENTION: APOLIPOPROTEIN A-I AGONISTS
; TITLE OF INVENTION: AND THEIR USE TO TREAT DYSLIPIDEMIC DISORDERS
; NUMBER OF SEQUENCES: 258
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/453,826
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/940,095
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 009196-0004-999

```

; INFORMATION FOR SEQ ID NO: 73:
;
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 22 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear

```

SEQUENCE DESCRIPTION: SEQ ID NO: 73:
5-989-73

Query Match 27.9%; Score 56; DB 4; Length 22;
Best Local Similarity 42.9%; Pred. No. 0.58;
Matches 9; Conservative 5; Mismatches 7; Indels

QY 16 PVTQEFWDLNLEKETEGLRQEM 36
|||:|:|:|:
Db 1 PVLDEFWEKLNEXLEALKQKL 21

RESULT 14

US-09-453-834-73
; Sequence 73, Application US/09453834
; Patent No. 6753313
; GENERAL INFORMATION:
; APPLICANT: Dasseux, Jean-Louis
; APPLICANT: Sekul, Renate
; APPLICANT: Buttner, Klaus
; APPLICANT: Cornut, Isabelle
; APPLICANT: Metz, Gunther
; APPLICANT: Dufourcq, Jean
; TITLE OF INVENTION: APOLIPOPROTEIN A-I AGONISTS
; NUMBER OF SEQUENCES: 258
; NUMBER OF INVENTION: AND THEIR USE TO TREAT DYSLIPIDEMIC DISORDERS
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/453,834
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/940,095
; FILING DATE: 29-SEP-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 009196-0004-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-493-4935
; TELEFAX: 650-493-5556
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 73:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: NO. 6753313e
; FEATURE:
; NAME/KEY: Other
; LOCATION: 13
; OTHER INFORMATION: Xaa = Aib
US-09-453-834-73

Query Match 27.9%; Score 56; DB 4; Length 22;
Best Local Similarity 42.9%; Pred. No. 0.58;
Matches 9; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 16 PVTQEFWDLNLEKETEGLRQEM 36
|||:|:|:|:
Db 1 PVLDEFWEKLNEXLEALKQKL 21

RESULT 15

US-07-928-930A-8
; Sequence 8, Application US/07928930A
; Patent No. 5344822

; GENERAL INFORMATION:
; APPLICANT: Levine, Daniel M.; Parker, Thomas S.;
; APPLICANT: Rubin, Albert L.
; TITLE OF INVENTION: Methods Useful in Endotoxin
; TITLE OF INVENTION: Prophylaxis and Therapy
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/928,930A
; FILING DATE: 19920812
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5344822man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: ROGO 206
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884

; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear

US-07-928-930A-8

Query Match 25.4%; Score 51; DB 1; Length 22;
Best Local Similarity 47.6%; Pred. No. 2.7;
Matches 10; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 16 PVTQEFWDLNLEKETEGLRQEM 36
|||:|:|:|:
Db 1 PVLDEFREKLNLEALEALKQKM 21

Search completed: December 21, 2004, 12:15:07
Job time : 40 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2004, 11:46:06 ; Search time 38 Seconds
(without alignments)
283.586 Million cell updates/sec

Title: US-09-803-918A-2_COPY_156_267
Perfect score: 556
Sequence: 1 QXHLQELKSLPGEMRDR.....SPKVSFLSALBEYTKKLNQ 112

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 46225

Minimum DB seq length: 0
Maximum DB seq length: 112

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	ID	Description
1	67	12.1	71	IS1705	stathmin - African
2	66.5	12.0	104	A32731	somatoliberin prec
3	64	11.5	88	H87560	conserved hypotet
4	62.5	11.2	109	IWBO	ATPase inhibitor p
5	62	11.2	97	T46234	hypothetical prote
6	62	11.2	107	S76457	hypothetical prote
7	61	11.0	104	A12730	transcription regu
8	61	11.0	108	A59010	antifreeze protein
9	59.5	10.7	92	AB1439	small heat shock p
10	59.5	10.7	102	A13243	transcriptional re
11	59.5	10.7	102	T30423	tram protein - Agr
12	59	10.6	73	B70542	hypothetical prote
13	59	10.6	110	E81366	probable membrane
14	59	10.6	111	T50084	hypothetical coile
15	58.5	10.5	107	F69228	hypothetical prote
16	57.5	10.3	108	G83073	hypothetical prote
17	57	10.3	86	F64527	hypothetical prote
18	57	10.3	94	PS0349	myosin heavy chain
19	57	10.3	96	C98046	30S ribosomal prot
20	57	10.3	96	D95179	ribosomal protein
21	57	10.3	110	B83481	hypothetical prote
22	56.5	10.2	91	I54248	apolipoprotein A-I
23	56.5	10.2	99	JU0038	nonhistone chromos
24	56.5	10.2	104	E69805	hypothetical prote
25	56	10.1	79	F91093	type III secretion
26	56	10.1	93	AE1522	hypothetical prote
27	56	10.1	98	E85747	hypothetical prote
28	56	10.1	101	C83539	morphogene protein
29	55.5	10.0	97	A11081	a small heat shock

30	55.5	10.0	98	2	S44095	intermediate filam
31	55.5	10.0	100	2	T50611	hypothetical prote
32	55.5	10.0	109	2	T29123	hypothetical prote
33	55	9.9	97	2	AF1306	conserved hypotet
34	55	9.9	97	2	AF1678	conserved hypotet
35	55	9.9	101	2	AC0895	conserved hypotet
36	55	9.9	109	2	F83736	transcription regu
37	54.5	9.8	88	2	I48251	apolipoprotein C1
38	54.5	9.8	106	2	B87263	conserved hypotet
39	54.5	9.8	107	2	H90509	hypothetical prote
40	54	9.7	67	2	T05224	hypothetical prote
41	54	9.7	83	2	A33854	outer membrane lip
42	53.5	9.6	93	2	AE0923	peptidyl-prolyl ci
43	53.5	9.6	97	2	D89792	conserved hypotet
44	53.5	9.6	112	2	T39864	hypothetical prote
45	53.5	9.6	112	2	A72717	hypothetical prote

ALIGNMENTS

RESULT 1

IS1705
stathmin - African clawed frog (fragment)
C:Species: Xenopus laevis (African clawed frog)
C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004
C:Accession: IS1705
R:Maucuer, A.; Moreau, J.; Mechali, M.; Sobel, A.
J. Biol. Chem. 268, 16420-16429, 1993
A:Title: Stathmin gene family: phylogenetic conservation and developmental regulation i
A:Reference number: A47345; MUID:93346387; PMID:8344928
A:Accession: IS1705
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-71 <MAU>
A:Cross-references: UNIPROT:Q09005; EMBL:X71432; NID:g937171; PIDN:CAA50563.1; PID:g937
C:Genetics:
A:Gene: XO20
C:Superfamily: stathmin

Query Match 12.1%; Score 67; DB 2; Length 71;
Best Local Similarity 38.5%; Pred. No. 53;
Matches 20; Conservative 11; Mismatches 13; Indels 8; Gaps 3;

QY 35 YSEDLRQLAARLEALKENGGAARLAAYHAKATEHLSTLSEKAKPALEDLRQG 86
Db 19 FSKMAEKLTTKWEAIKENREAQMA---AK----LERLEKDK-KLEEIRKG 62

RESULT 2

A32731
somatoliberin precursor - rat
N:Alternate names: growth hormone-releasing hormone
C:Species: Rattus norvegicus (Norway rat)
C>Date: 13-Jul-1990 #sequence_revision 13-Jul-1990 #text_change 09-Jul-2004
C:Accession: A32731; A41366; I67421
R:Mayo, K.E.; Cerelli, G.M.; Rosenfeld, M.G.; Evans, R.M.
Nature 314, 464-467, 1985
A:Title: Characterization of cDNA and genomic clones encoding the precursor to rat hypo
A:Reference number: A32731; MUID:85163768; PMID:3920534
A:Accession: A32731
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-104 <MAY>
A:Cross-references: UNIPROT:P09916; GB:X02320
R:Gonzalez-Crespo, S.; Boronat, A.
Proc. Natl. Acad. Sci. U.S.A. 88, 8749-8753, 1991
A:Title: Expression of the rat growth hormone-releasing hormone gene in placenta is dir
A:Reference number: A41366; MUID:92020929; PMID:1924334
A:Accession: A41366
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-104 <GON>

A:Cross-references: GB:M73486; NID:G204311; PIDN:AAA41220.1; PID:G204312
R:Srivastava, C.H.; Monts, B.S.; Rothrock, J.K.; Peredo, M.J.; Pescovitz, O.H.
Endocrinology 136, 1502-1508, 1995
A:Title: Presence of a spermatogenic-specific promoter in the rat growth hormone-release
A:Reference number: I53290; MUID:95203210; PMID:7895659
A:Accession: I67421
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-104 <RES>
A:Cross-references: EMBL:U10156; NID:G498584; PIDN:AAC52184.1; PID:G498585
C:Genetics:
A:Gene: GHRH
C:Superfamily: glucagon
C:Keywords: duplication

Query Match 12.0%; Score 66.5; DB 2; Length 104;
Best Local Similarity 32.9%; Pred. No. 86;
Matches 23; Conservative 10; Mismatches 32; Indels 5; Gaps 3;

QY 20 RARAHVDALRTHLAPYSDELQRLAARL--EALKENGGRARLAHYHAKATEHLSTL-SEKA 76
Db 27 RVRRHADAFT--SSYRRLGQYARKLLHETMNRQGRNORSRFRNRLDRVWAEDK 84

QY 77 KPALEDLROG 86
Db 85 QVALESILQG 94

RESULT 3
H87560
conserved hypothetical protein CC2513 [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C:Accession: H87560
R:Niezman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: AB7249; MUID:21173698; PMID:11259647
A:Accession: H87560
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-88 <STO>
A:Cross-references: UNIPROT:Q9A5D7; GB:AE005673; NID:G13424070; PIDN:AAK24484.1; GSPDB:G
C:Genetics:
A:Gene: CC2513

Query Match 11.5%; Score 64; DB 2; Length 88;
Best Local Similarity 32.7%; Pred. No. 1.2e+02;
Matches 18; Conservative 11; Mismatches 22; Indels 4; Gaps 1;

QY 20 RARAHVDALRTHLAPYSDELQRLAARL--EALKENGGRARLAHYHAKATEHLSTLSE 74
Db 9 RALADJTAIRDYIGQFSPLAAQRMALRLKTAADS----LAEPYPERGLATATLRE 59

RESULT 4
IMBO
ATPase inhibitor precursor, mitochondrial - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 19-Feb-1984 #sequence_revision 28-Apr-1995 #text_change 09-Jul-2004
C:Accession: C27382; A01336
R:Walker, J.E.; Gay, N.J.; Powell, S.J.; Kostina, M.; Dyer, M.R.
Biochemistry 26, 8613-8619, 1987
A:Title: ATP synthase from bovine mitochondria: sequences of imported precursors of olig
A:Reference number: A90527; MUID:88163536; PMID:2894843
A:Accession: C27382
A:Molecule type: mRNA
A:Residues: 1-109 <WAL>
A:Cross-references: UNIPROT:P01096; GB:M22559; NID:G162712; PIDN:AAA30396.1; PID:G162713
R:Frangione, B.; Rosenwasser, E.; Penefsky, H.S.; Pullman, M.E.
Proc. Natl. Acad. Sci. U.S.A. 78, 7403-7407, 1981

A:Title: Amino acid sequence of the protein inhibitor of mitochondrial adenosine triphos
A:Reference number: A01336; MUID:82150878; PMID:6461003
A:Accession: A01336
A:Molecule type: protein
A:Residues: 26-54,'Q',56-109 <FRA>
A:Note: there are several internal duplications near the carboxyl end (residues 73-97)
C:Comment: This peptide is thought to be a regulatory component of the ATP-synthesizing
C:Superfamily: ATPase inhibitor, mitochondrial
C:Keywords: ATPase inhibitor; mitochondrial
F:1-25/Domain: transit peptide (mitochondrion) #status predicted <SIG>
F:26-109/Product: ATPase inhibitor #status experimental <MAT>

Query Match 11.2%; Score 62.5; DB 1; Length 109;
Best Local Similarity 27.5%; Pred. No. 1.9e+02;
Matches 22; Conservative 13; Mismatches 36; Indels 9; Gaps 2;

QY 14 GBEMDRARAHVDALR-----THLAPYSDELQRLAARL--EALKENGGRARLAHYHAKA 65
Db 26 GSESGDNVRSAGAVRDAGGAFGKREQAEEERYFRARAKEQLAALKKHENEIS-HHAK 84

QY 66 TEHLSTLSEKAKPALEDLRQ 85
Db 85 IERLQKEIERHKQSIKLLQ 104

RESULT 5
T46234
hypothetical protein T9C5.160 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
C:Accession: T46234
R:Rieger, M.; Gabel, C.; Mueller-Auer, S.; Schaefer, M.; Zipp, M.; Mewes, H.W.; Lemcke,
submitted to the Protein Sequence database, December 1999
A:Reference number: Z23026
A:Accession: T46234
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-97 <RIE>
A:Cross-references: UNIPROT:Q9SCK2; EMBL:AL132964
A:Experimental source: cultivar Columbia; BAC clone T9C5
C:Genetics:
A:Map position: 3
A:Note: T9C5.160

Query Match 11.2%; Score 62; DB 2; Length 97;
Best Local Similarity 21.0%; Pred. No. 1.8e+02;
Matches 17; Conservative 22; Mismatches 40; Indels 2; Gaps 1;

QY 1 OKLHELQEKLSPLGEMDRARAHVDALRTHLAPYSDELQRLAARL--EALKENGGRARLA 60
Db 13 EVELSRRRNGELEREEMEMKEMVQLWRRTVV--AEAEERLCSQLAELESVLDQARD 70

QY 61 YHAKATEHLSTLSEKAKPALE 81
Db 71 YHSRIVFLMDQISRLSSSSLE 91

RESULT 6
S76457
hypothetical protein - Synechocystis sp. (strain PCC 6803)
C:Species: Synechocystis sp.
A:Variety: PCC 6803
C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C:Accession: S76457
R:Kansko, T.; Sato, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasud
o, K.; Okumura, S.; Shimpo, S.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
s.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S76457
A>Status: preliminary
A:Molecule type: DNA

A;Residues: 1-107 <KAN>
A;Cross-references: UNIPROT:P74485; EMBL:D90915; GB:AB001339; NID:g1653604; PIDN:BAA1858
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 11.2%; Score 62; DB 2; Length 107;
Best Local Similarity 23.9%; Pred.No. 2e+02;
Matches 21; Conservative 17; Mismatches 30; Indels 20; Gaps 3;

Qy 7 QEKLSPLGEEMRDRARAVHDALRTHLPYSDLRQRILA-----ARLEALKENGGA RLA 59
||| : : : : : : : : : : : : : : : : : :
Db 15 QEK-----EAYKEKVRAKIKDKNAQIDMTAEAREKAADANVYQKSLDQLQAORDALMG 69
||| : : : : : : : : : : : : : : : : : :

Qy 60 EYHAKATEHLSTLSEKAPPALEDRLQGL 87
:|| : : : : : : : : : : : : : : : : : |
Db 70 KWH-----DLQQSGEA AWEELQAGL 89
||| : : : : : : : : : : : : : : : : : |

RESULT 7
AI2730
transcription regulator, Arsr family Atul254 [imported] - Agrobacterium tumefaciens (strain C58)
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C;Accession: AI2730
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.; erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kuttyavin, T.; Levy, R.; Li, M.; McClellan, P.; Karp, P.; Romero, F.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AI2730
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-104 <KUR>
A;Cross-references: UNIPROT:Q8UPZ0; GB:AB008688; PIDN:AAL42263.1; PID:g17739660; GSPDB:G
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: Atul254
A;Map position: circular chromosome

Query Match 11.0%; Score 61; DB 2; Length 104;
Best Local Similarity 30.8%; Pred.No. 2.4e+02;
Matches 20; Conservative 10; Mismatches 23; Indels 12; Gaps 3;

Qy 33 APYSDELQRQLAARLEALKENGGA RLA EYHAKATEHLSTLSEKAPPALEDRLR---CGLLP 89
||| : : : : : : : : : : : : : : : : : |
Db 46 APCDDMPRTSLSSHWRILREAG---LIRTHKGQVENISVV-----RLEDWEARFFGLLP 96
||| : : : : : : : : : : : : : : : : : |

Qy 90 VLESF 94
: |
Db 97 SILKF 101
: |

RESULT 8
A59010
antifreeze protein LG-12 - longhorn sculpin
C;Species: Myoxocephalus octodecimpinosus (longhorn sculpin)
C;Date: 10-Oct-1997 #sequence_revision 10-Oct-1997 #text_change 09-Jul-2004
C;Accession: A59010
R;Deng, G.; Andrews, D.W.; Laursen, R.A.
FEBS Lett. 402, 17-20, 1997
A;Reference number: A59010; MUID:97165956; PMID:9013849
A;Accession: A59010
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-108 <LAU>
A;Cross-references: UNIPROT:p80961
A;Experimental source: blood plasma
C;Keywords: antifreeze; plasma; pyroglutamic acid
F;/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 11.0%; Score 61; DB 2; Length 108;

```
Best Local Similarity   22.9%; Pred. No. 2.5e+02;
Matches    25; Conservative    24; Mismatches    42; Indels     18; Gaps      4;

Qy       10 LSPGLGEMRRARAHDVALRTHLAPYSDEL--QRLAARLEALKENGARGALAEYHAKATE 67
Db            |::| : ::| : ||::|| : |::| : |:| :|::| :|::| :|:| :|:| :|:~
               -----YEEMTKLIQDWTETIIRSODLANQAQAFVEDKKTQLPVAIQIE 57
Qy       68 HLSTLS-----EKAKPALEDLRQGILLPVLESFKVSFLSALBEYYTCKLNQTQ 112
Db            |::| :|:| :|::| :|::| :|::| :|::| :|::| :|::| :|::| :|::| :|:~
               QMKTATVNVEEQIRPLTANVQHLPQIDNfq---QMEEAIKKLTDDQ 102

RESULT 9
AB1439 small heat shock protein of Clostridium acetobutylicum homolog lin0049 [imported] - Lis-
C:Species: Listeria innocua
C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: AB1439
F.;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecke
r.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Faihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mo
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Soland, J.A.; Voss, H.; Wehlend
A.:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; PMID:21537279; PMID:11679669
A:Accession: AB1439
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-97 <GUI>
A:Cross-references: UNIPROT:Q92FQ4; GB:AL592022; PIDN:CAC95282.1; PID:g16412470; GSPPDB:X
A:Experimental source: strain Clip11262
C:Genetics:
A:Gene: lin0049

Query Match          10.7% ; Score 59.5; DB 2; Length 97;
Best Local Similarity 25.3%; Pred. No. 2.9e+02;
Matches    24; Conservative    20; Mismatches    30; Indels     21; Gaps      5;

Qy       7 QEKLSPGLEMRDRARAH-----VDALRTHLAPYSDELR-----QRLAARLEALKE 52
Db            |::| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:~
               4 QRIMSP--SELDRAKYTGSGRDIEDIIISRLSQDLRSWEGQAFRFDDQFEOLK- 60
Qy       53 NGGARLAELYHAKATEHLSTLSEXAKPAEIDLROGL 87
Db            |::| :|:| :|:~
               61 ---PKYTEF-A NLMDQINDOLEXTANA VEHDQOL 91

RESULT 10
AI13243 transcriptional repressor traM [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C:Species: Agrobacterium tumefaciens
C>Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C:Accession: AI13243
F.;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo,
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kuttyavin, T.; Levy, R.; Li, M.; McClellan
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; PMID:21608550; PMID:11743193
A:Accession: AI13243
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-102 <CUR>
A:Cross-references: UNIPROT:Q44452; GB:AEO08690; PIDN:AAL46367.1; PID:g17744157; GSPPDB:X
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: traM
A:Genome: plasmid

Query Match          10.7% ; Score 59.5; DB 2; Length 102;
```

Query Match 11.0%; Score 61; DB 2; Length 108;

Query Match
10.7%; Score 59.5; DB 2; Length 102;

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2004, 11:47:05 ; Search time 186 Seconds
(without alignments)

346.462 Million cell updates/sec

Title: US-09-803-918A-2_COPY_156_267

Perfect score: 556

Sequence: 1 QKHLQKLSPLGEMRDR.....SPKVSFLSALEYTKKLNTQ 112

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 350892

Minimum DB seq length: 0

Maximum DB seq length: 112

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 02.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	327	58.8	67	Q9Y355	Q9Y355 homo sapien
2	246	44.2	50	Q8TDB0	Q8TDB0 homo sapien
3	109.5	19.7	102	Q98TJ9	Q98TJ9 platicthys
4	97	17.4	107	1 APE_SAI5C	Q28955 salmieri sci
5	93	16.7	107	1 APE_SACMU	Q28502 macaca mula
6	82	14.7	96	2 Q7P280	Q7P280 fusobacteri
7	79	14.2	96	2 Q8RH51	Q8RH51 fusobacteri
8	78.5	14.1	110	2 Q7NRA9	Q7NRA9 chromobacte
9	77	13.8	66	2 Q9DFQ4	Q9DFQ4 gillichthys
10	76.5	13.8	96	2 Q9T851	Q9T851 sus scrofa
11	72.5	13.0	104	2 Q7ZU66	Q7ZJ66 thermus the
12	72.5	13.0	104	2 AAS81257	AAS81257 thermus t
13	70	12.6	102	2 Q29258	Q29258 sus scrofa
14	68	12.2	85	2 Q7Q1M4	Q7Q1M4 anopheles g
15	67.5	12.1	79	2 Q6LD50	Q6LD50 mus sp. apo
16	67.5	12.1	79	2 AAB35539	AAB35539 mus sp. a
17	67	12.1	71	1 STIB_XENLA	Q09005 xenopus lae
18	66.5	12.0	104	1 SLIB_RAT	P09916 rattus norv
19	66.5	12.0	105	2 Q8ZV19	Q8ZV19 pyrobaculum
20	66	11.9	84	2 Q82YMO	Q82YMO enterococcu
21	65	11.7	84	2 Q7U6M0	Q7U6M0 synecococc
22	64.5	11.6	98	2 P91785	P91785 onchocerca
23	64	11.5	88	2 Q9A5D7	Q9A5D7 caulobacter
24	62.5	11.2	109	1 IATP_BOVIN	P01096 bos taurus
25	62	11.2	97	2 Q9SCX2	Q9SCX2 arabidopsis
26	62	11.2	107	2 P74485	P74485 synecocyst
27	61.5	11.1	76	2 Q8FRA8	Q8FRA8 corynebacte
28	61	11.0	101	2 Q6Y7N2	Q6Y7N2 staphylococ
29	61	11.0	101	2 AA047534	AA047534 staphyloc
30	61	11.0	104	2 Q8UF20	Q8UF20 agrobacteri
31	61	11.0	109	2 Q84GY9	Q84GY9 photorhabdu

32	60.5	10.9	83	2	Q711L8	Q711L8 fasciola he
33	60.5	10.9	83	2	CAC86373	CAC86373 fasciola
34	60.5	10.9	86	2	Q8P6R7	Q8P6R7 xanthomonas
35	60.5	10.9	89	2	Q8LQ79	Q8LQ79 oryza sativ
36	60.5	10.9	104	2	Q8R814	Q8R814 thermoaer
37	60	10.8	85	2	Q9QUL5	Q9QUL5 human immun
38	59.5	10.7	88	1	APC1_TUPGL	Q9X8N5 tupaia glis
39	59.5	10.7	97	2	Q92FQ4	Q92FQ4 listeria in
40	59.5	10.7	102	1	TRAM_AGR75	Q4452 agrobacteri
41	59.5	10.7	102	2	Q9S6E6	Q9S6E6 agrobacteri
42	59.5	10.7	107	2	Q35693	Q35693 mus musculu
43	59.5	10.7	110	2	Q6PRF2	Q6PRF2 helicobacte
44	59.5	10.7	110	2	Q8KKD9	Q8KKD9 helicobacte
45	59.5	10.7	110	2	AAS93839	AAS93839 helicobac

ALIGNMENTS

RESULT 1
Q9Y355 PRELIMINARY; PRT; 67 AA.
AC Q9Y355; PRELIMINARY; PRT; 67 AA.
DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Apolipoprotein A1 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99216322; PubMed=10198255;
RA Hamidi Asl K., Liepnieks J.J., Nakamura M., Parker F., Benson M.D.;
RT "A novel apolipoprotein A-1 variant, Arg173Pro, associated with
cardiac and cutaneous amyloidosis.";
RL Biochem. Biophys. Res. Commun. 257:584-588(1999).
DR EMBL; AF148963; AAD34604.1; -
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0008289; F:lipid binding; IEA.
DR GO; GO:0008869; P:lipid transport; IEA.
DR GO; GO:0042157; P:lipoprotein metabolism; IEA.
DR InterPro; IPR000074; Apolipo A.E.C3.
DR Pfam; PF01442; Apolipoprotein_1.
KW Lipoprotein.
FT NON_TER 1 1
FT NON_TER 67 67
SQ SEQUENCE 67 AA; 7433 MW; 525E1FEB7BDD5AFB CRC64;

Query Match 58.8%; Score 327; DB 2; Length 67;
Best Local Similarity 98.5%; Pred.No. 28-19;
Matches 66; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	10	LSPLSEWRDRARAHVDALRTHLAPYSDELRLQRLAARLEALKENGARLAAYHAKATEHL	69
Db	1	LSPLSEWRDRARAHVDALRTHLAPYSDELRLQRLAARLEALKENGARLAAYHAKATEHL	60
QY	70	STLSEKA 76	
Db	61	STLSEKA 67	

RESULT 2

Q8TDB0 PRELIMINARY; PRT; 50 AA.
ID Q8TDB0; PRELIMINARY; PRT; 50 AA.
AC Q8TDB0;
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Apolipoprotein A-1 A175P variant (Fragment).
OS Homo sapiens (Human).


```

RESULT 5
APE_MACMU STANDARD; PRT; 107 AA.
ID -- APE MACMU
AC Q28502;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Apolipoprotein E (Apo-E) (Fragment).
GN Name=APOE;
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Macaca.
OC NCBI_TaxID=9544;
RN [1]
SEQUENCE FROM N.A.
MEDLINE=96225955; PubMed=8635577;
RA Morelli L., Wei L., Amorim A., McDermid J., Abee C.R., Frangione B.,
RA Walker L.C., Levy E.;
RT "Cerebrovascular amyloidosis in squirrel monkeys and rhesus monkeys:
RT apolipoprotein E genotype."
RL FEBS Lett. 379:132-134 (1996).
CC -!- FUNCTION: Mediates the binding, internalization, and catabolism of
CC lipoprotein particles. It can serve as a ligand for the LDL (apo
CC B/E) receptor and for the specific apo-E receptor (chylomicron
CC remnant) of hepatic tissues.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Secreted in plasma.
CC -!- SIMILARITY: Belongs to the apolipoprotein A1/A4/E family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U52030; AAC50441.1; --
CC HSSP; P02649; INFN.
DR InterPro; IPR000074; Apolipoproteins.
DR InterPro; IPR009074; Apolipo_A_E_C3.
DR Pfam; PF01442; Apolipoprotein; 1.
DR Chylomicron; Glycoprotein; HDL; Heparin-binding; Lipid transport;
KW Plasma; Repeat; VLDL.
FT NON TER 1
FT DOMAIN 89 99 LDL receptor binding (Potential).
FT DOMAIN 93 96 Heparin-binding (by similarity).
FT DOMAIN 11 >107 8 X 22 AA approximate tandem repeats.
FT REPEAT 11 32 1.
FT REPEAT 33 54 2.
FT REPEAT 55 76 3.
FT REPEAT 77 98 4.
FT REPEAT 99 >107 5.
FT NON TER 107 107
SQ SEQUENCE 107 AA; EID38C32F5AACB23 CRC64;
Query Match 16.7%; Score 93; DB 1; Length 107;
Best Local Similarity 22.1%; Pred. No. 3.8;
Matches 23; Conservative 27; Mismatches 54; Indels 0; Gaps 0;
Qy 5 ELQKSLPLGEEMDRADRAHVDAIRLTHLAPYSDELRLQRLARLEALKENGCGARLAAYHAK 64
Db 4 QVTQELTLMDETMKELKAYKSELEEQLSVAETRLARLSKELQAQARLGADMDVRSR 63
Qy 65 ATEHLSLTSKAKPALDELRLGGLLPVLSEFKVPSFLSALEYTKK 108
Db 64 LVQYRSEVQAMLGQSTBELRLARLASHLRKLRKRLRLDADDLQKR 107
RESULT 6
Q7P280 PRELIMINARY; PRT; 96 AA.
ID -- Q7P280

```

```

QY 2 KLHELOEKLSPGEMDRARAHVDALRTHLAPYSDELQRQLAARLEALKENGARLAAY 61
DB 10 KAEELGVATDKAKELKDKTKVAKAEELKNKVKVYDKAKELKEGAKGASL 69
QY 62 HAKATEHLSTLSKAKPALEDL 83
DB 70 KDKAAKAEELKDKITDGADSL 91

RESULT 8
Q7NRA9
ID Q7NRA9 PRELIMINARY; PRT; 110 AA.
AC Q7NRA9;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DE Hypothetical protein.
GN OrderedLocusNames=CV3874;
OS Chromobacterium violaceum.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Chromobacterium.
OX NCBI_TaxID=536;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 12472 / DSM 30191;
RX MEDLINE=22892880; PubMed=14500782;
RA Vasconcelos A.T.R., de Almeida D.F., Hungria M., Guimaraes C.T.,
RA Antonio R.V., Almeida F.C., de Almeida L.G.P., de Almeida R.,
RA Alves-Gomes J.A., Andrade E.M., Araripe J., de Araujo M.F.P.,
RA Batolfi-Filho S., Azevedo V., Baptista A.J., Bataus L.A.M.,
RA Batista J.S., Belo A., van den Berg C., Bogo M., Bonatto S.,
RA Bordignon J., Brighido M.M., Brito C.A., Brocchi M., Burity H.A.,
RA Camargo A.A., Cardoso D.P., Carneiro N.P., Carraro D.M.,
RA Carvalho C.M.B., Cascado J.C.M., Cavada B.S., Chueire L.M.O.,
RA Creczynski-Pasa T.B., Cunha-Junior N.C., Fagundes N., Falcao C.L.,
RA Pantinatti F., Farias I.P., Felipe M.S.S., Ferrari L.P., Ferro J.A.,
RA Ferro M.I.T., Franco G.R., Freitas N.S.A., Furlan L.R.,
RA Gazzinelli R.T., Gomes E.A., Goncalves P.R., Grangeiro T.B.,
RA Grattapaglia D., Grisar E.C., Hanna E.S., Jardim S.N., Laurino J.,
RA Leoi L.C.T., Lima L.F.A., Loureiro M.F., Lyra M.C.C.P.,
RA Madeira H.M.F., Manfio G.P., Maranhao A.Q., Martins W.S.,
RA di Mauro S.M.Z., de Medeiros R.B., Meisner R.V., Moreira M.A.M.,
RA Nascimento F.F., Nicolas M.F., Oliveira J.G., Oliveira S.C.,
RA Faixao R.P.C., Parente J.A., Pedrosa F.O., Pena S.D.J., Pereira J.O.,
RA Pereira M., Pinto L.S.R.C., Pinto L.S., Porto J.I.R., Potrich D.P.,
RA Ramalho-Neto C.E., Reis A.M.M., Rigo L.U., Rondinelli E.,
RA Santos E.B.P., Santos F.R., Schneider M.P.C., Seunanez H.N.,
RA Silva A.M.R., da Silva A.L.C., Silva D.W., Silva R., Simoes I.C.,
RA Simon D., Soares C.M.A., Soares R.B.A., Souza E.M., Souza K.R.L.,
RA Souza R.C., Steffens M.B.R., Steindel M., Teixeira S.R., Urmenyi T.,
RA Vettore A., Wassem R., Zaha A., Simpson A.J.G.;
RT "The complete genome sequence of Chromobacterium violaceum reveals
RT remarkable and exploitable bacterial adaptability.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:11660-11665(2003).
DR EMBL; AE016923; AAQ61536.1; --
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 110 AA; 12757 MW; 6B6CC3929C615493 CRC64;

Query Match 14.1%; Score 78.5; DB 2; Length 110;
Best Local Similarity 32.6%; Pred. No. 59;
Matches 31; Conservative 12; Mismatches 31; Indels 21; Gaps 4;

QY 3 LHELOEKLSPGSE-----MRDRAR-----AHVDALRTHLA-----PYSDELQR 42
DB 10 LAELIEQLPTELLLEAANLRPRFSSLYGRSEAHVQQLKTLRQEGRDQLSDQREA 69
QY 43 LAARLEALKENGARLAAYHAKATEHLSTLSEKAK 77
DB 70 L-HRVLIIVREETQRQLANWAGQVKDELTLTKSSK 103

RESULT 9
QY 1 OKLHELOEKLSPGEMDR--ARAHVDALRTHLAPYSDELQRQLAARLEALKENGGA 56

```

```

Q9DFQ4
ID Q9DFQ4 PRELIMINARY; PRT; 66 AA.
AC Q9DFQ4;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Apolipoprotein A-1 (Fragment)
OS Gillichthys seta (Shortjaw muddeucker).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Gobioidae;
OC Gobiidae; Gillichthys.
OX NCBI_TaxID=79683;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=21117151; PubMed=11172064;
RA Gracey A.Y., Troll J.V., Somero G.N.;
RT "Hypoxia-induced gene expression profiling in the euryoxic fish
RT Gillichthys mirabilis.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:1993-1998(2001).
DR EMBL; AF266178; AAG13298.1; --
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0008289; P:lipid binding; IEA.
DR GO; GO:0006869; P:lipid transport; IEA.
DR GO; GO:0042157; P:lipoprotein metabolism; IEA.
DR InterPro; IPR000074; Apolipoprotein.
DR InterPro; IPR009074; Apolipo_A_E_C3.
DR Pfam; PF01442; Apolipoprotein; 1.
KW Lipoprotein.
FT NON_TER 1
FT NON_TER 66
SQ SEQUENCE 66 AA; 7637 MW; 0B5B673FA8C91B7D CRC64;

Query Match 13.8%; Score 77; DB 2; Length 66;
Best Local Similarity 33.3%; Pred. No. 47;
Matches 14; Conservative 13; Mismatches 15; Indels 0; Gaps 0;

QY 9 KLSPGEMDRARAHVDALRTHLAPYSDELQRQLAARLEAL 50
DB 15 KLEPTEEMERAKIEVNBETKSLVPIIDTVREKITERVKAV 56

RESULT 10
Q9TS51
ID Q9TS51 PRELIMINARY; PRT; 96 AA.
AC Q9TS51;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Myosin heavy chain (fragment).
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE.
RX MEDLINE=95246271; PubMed=7537184;
RA Inonata T., Hanawa H., Miyaniishi T., Yajima E., Nakayama S., Maita T.,
RA Kodama M., Izumi T., Shibata A., Abo T.;
RT "Localization of porcine cardiac myosin epitopes that induce
RT experimental autoimmune myocarditis.";
RL Circ. Res. 76:726-733(1995).
DR InterPro; IPR002928; Myosin tail.
DR Pfam; PF01576; Myosin tail 1; 1.
SQ SEQUENCE 96 AA; 11137 MW; 69943C386226BEEC CRC64;

Query Match 13.8%; Score 76.5; DB 2; Length 96;
Best Local Similarity 37.9%; Pred. No. 75;
Matches 22; Conservative 11; Mismatches 16; Indels 9; Gaps 2;

QY 1 OKLHELOEKLSPGEMDR--ARAHVDALRTHLAPYSDELQRQLAARLEALKENGGA 56

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2004, 11:55:03 ; Search time 152 Seconds
(without alignments)
264.327 Million cell updates/sec

Title: US-09-803-918A-2_COPY_156_267
Perfect score: 556
Sequence: 1 QKLUHEIQELSPGEMRDR.....SFKVSFLSALAEYTKKLTNQ 112

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 1169828

Minimum DB seq length: 0
Maximum DB seq length: 112

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:.*
1: Geneseqp1980s:.*
2: Geneseqp1990s:.*
3: Geneseqp2000s:.*
4: Geneseqp2001s:.*
5: Geneseqp2002s:.*
6: Geneseqp2003as:.*
7: Geneseqp2003bs:.*
8: Geneseqp2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	244	43.9	49	6	ABR44037 Human Apo
2	244	43.9	49	6	ABR44039 Human Apo
3	244	43.9	49	6	ABR44038 Human Apo
4	240	43.2	49	6	ABR44040 Baboon Ap
5	186	33.5	49	6	ABR44041 Rabbit Ap
6	161	29.0	33	7	ADC29656 Antioxida
7	157.5	28.3	48	6	ABR44042 Pig ApoA-
8	153	27.5	33	7	ADC29655 Antioxida
9	129	23.2	49	6	ABR44044 Chicken A
10	113	20.3	47	6	ABR44043 Rat ApoA-
11	111	20.0	85	8	ADP87442 Human apo
12	100	18.0	38	4	AAO08726 Human pol
13	100	18.0	85	8	ADP87441 Monkey ap
14	99	17.8	85	8	ADP87439 Pig apoli
15	96	17.3	85	8	ADP87440 Baboon ap
16	94	16.9	18	7	ADC29635 Antioxida
17	91	16.4	22	4	AGS2608 Apolipop
18	90	16.2	18	7	ADC29638 Antioxida
19	90	16.2	96	8	ADP87420 Squirrel
20	87	15.6	18	7	ADC29642 Antioxida
21	87	15.6	96	8	ADP87422 Baboon ap
22	87	15.6	96	8	ADP87423 Rhesus ma
23	87	15.6	96	8	ADP87421 Monkey ap
24	86	15.5	18	7	ADC29636 Antioxida
25	86	15.5	18	7	ADC29639 Antioxida

26	86	15.5	18	7	ADC29634	Adc29634 Antioxida
27	85	15.3	18	7	ADC29637	Adc29637 Antioxida
28	84.5	15.2	96	8	ADP87430	Adp87430 Gorilla a
29	84	15.1	18	7	ADC29627	Adc29627 Antioxida
30	83	14.9	18	7	ADC29640	Adc29640 Antioxida
31	83	14.9	18	7	ADC29651	Adc29651 Antioxida
32	82	14.7	17	2	AAR98474	Aar98474 Anti-arte
33	82	14.7	18	7	ADC29649	Adc29649 Antioxida
34	82	14.7	18	7	ADC29646	Adc29646 Antioxida
35	82	14.7	18	7	ADC29644	Adc29644 Antioxida
36	82	14.7	18	7	ADC29647	Adc29647 Antioxida
37	82	14.7	18	7	ADC29650	Adc29650 Antioxida
38	81	14.6	17	6	ABR44036	AbR44036 Human Apo
39	81	14.6	17	6	ABR44035	AbR44035 Human Apo
40	81	14.6	96	8	ADP87428	Adp87428 Orangutan
41	81	14.6	96	8	ADP87429	Adp87429 Chimpanze
42	81	14.6	110	2	AAR12387	Aar12387 Amplified
43	80.5	14.5	107	8	ADP87437	Adp87437 Mouse apo
44	80.5	14.5	107	8	ADP87436	Adp87436 Mouse apo
45	80.5	14.5	107	8	ADP87434	Adp87434 Mouse apo

ALIGNMENTS

RESULT 1
ABR44037
ID ABR44037 standard; peptide; 49 AA.
XX ABR44037;
XX
DT 04-AUG-2003 (first entry)
XX
DE Human ApoA-I protein fragment.
XX
KW Apolipoprotein A-I; ApoA-I; cardiant; gene therapy; human;
KW cardiovascular disorder.
XX
OS Homo sapiens.
XX
PN WO2003035691-A1.
XX
PD 01-MAY-2003.
XX
PF 25-OCT-2002; 2002WO-CA001615.
XX
PR 26-OCT-2001; 2001US-0335075P.
XX
PA (XENO-) XENON GENETICS INC.
PA (UYBR-) UNIV BRITISH COLUMBIA.
PI Brownlie AJ, Dube M, Samuels M, Hayden MR;
XX
DR WPI; 2003-449258/42.
XX
PT New Apolipoprotein A-1 mutation, useful for preparing a composition for
PT treating cardiovascular disorder.
XX
PS Disclosure; Fig 4; 52pp; English.
XX
CC The invention relates to a mutation of the apolipoprotein A-I (ApoA-I)
CC comprising a leucine to proline variant in position 178. The mutant ApoA-
CC I polynucleotide is useful for preparing a composition for treating
CC cardiovascular disorder. Sequences ABR44037-44 represent ApoA-I fragments
CC from various species where the leucine at position 178 is highly
CC conserved except in pig (ABR44042)

Sequence 49 AA;

Query Match 43.9%; Score 244; DB 6; Length 49;
Best Local Similarity 100.0%; Pred. No. 1.9e-17;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 26 DALRTHLAPYSDELRLQRLAARLEALKENGARLAAYHAKATEHLSTLSE 74
Db 1 DALRTHLAPYSDELRLQRLAARLEALKENGARLAAYHAKATEHLSTLSE 49

RESULT 2
ABR44039
ID ABR44039 standard; peptide; 49 AA.
AC ABR44039;
DT 04-AUG-2003 (first entry)
DE Human ApoA-I wild-type protein fragment.
XX
KW Apolipoprotein A-I; ApoA-I; cardiant; gene therapy; human;
KW cardiovascular disorder.
XX
OS Homo sapiens.
XX
PN WO2003035691-A1.
XX
PD 01-MAY-2003.
XX
PF 25-OCT-2002; 2002WO-CA001615.
XX
PR 26-OCT-2001; 2001US-0335075P.
XX
PA (XENO-) XENON GENETICS INC.
XX
PY (UYBR-) UNIV BRITISH COLUMBIA.
XX
PI Brownlie AJ, Dube M, Samuels M, Hayden MR;
XX
PWPI; 2003-449258/42.
XX
DE New Apolipoprotein A-1 mutation, useful for preparing a composition for
PT treating cardiovascular disorder.
XX
PS Disclosure; Fig 4; 52pp; English.
XX
CC The invention relates to a mutation of the apolipoprotein A-I (ApoA-I)
CC comprising a leucine to proline variant in position 178. The mutant ApoA-
CC I polynucleotide is useful for preparing a composition for treating
CC cardiovascular disorder. Sequences ABR44037-44 represent ApoA-I fragments
CC from various species where the leucine at position 178 is highly
CC conserved except in pig (ABR44042)
XX
SQ Sequence 49 AA;
Query Match 43.9%; Score 244; DB 6; Length 49;
Best Local Similarity 100.0%; Pred. No. 1.9e-17;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 DALRTHLAPYSDELRLQRLAARLEALKENGARLAAYHAKATEHLSTLSE 74
Db 1 DALRTHLAPYSDELRLQRLAARLEALKENGARLAAYHAKATEHLSTLSE 49

RESULT 4
ABR44040
ID ABR44040 standard; peptide; 49 AA.
AC ABR44040;
DT 04-AUG-2003 (first entry)
DE Baboon ApoA-I protein fragment.
XX
KW Apolipoprotein A-I; ApoA-I; cardiant; gene therapy; baboon;
KW cardiovascular disorder.
XX
OS Papio sp.
XX
PN WO2003035691-A1.
XX
PD 01-MAY-2003.
XX
PF 25-OCT-2002; 2002WO-CA001615.
XX
PR 26-OCT-2001; 2001US-0335075P.
XX
PA (XENO-) XENON GENETICS INC.
XX
PY (UYBR-) UNIV BRITISH COLUMBIA.
XX
PI Brownlie AJ, Dube M, Samuels M, Hayden MR;
XX
PWPI; 2003-449258/42.
XX
DE New Apolipoprotein A-1 mutation, useful for preparing a composition for
PT treating cardiovascular disorder.
XX
PS Disclosure; Fig 4; 52pp; English.
XX

```

CC The invention relates to a mutation of the apolipoprotein A-I (ApoA-I) comprising a leucine to proline variant in position 178. The mutant ApoA-I polynucleotide is useful for preparing a composition for treating cardiovascular disorder. Sequences ABR44037-44 represent ApoA-I fragments from various species where the leucine at position 178 is highly conserved except in pig (ABR44042)

XX
SQ Sequence 49 AA;

Query Match 43.2%; Score 240; DB 6; Length 49;
Best Local Similarity 98.0%; Pred. No. 4.8e-17;
Matches 48; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 26 DALRTHLAPYSDELQRQLARLEALKENGCGARLAELYHAKATEHLSTLSE 74
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 DALRTHLAPYSDELQRQLARLEALKENGCGARLAELYHAKATEHLSTLSE 49

RESULT 5

ABR44041
ID ABR44041 standard; peptide; 49 AA.

AC ABR44041;

DT 04-AUG-2003 (first entry)

DE Rabbit ApoA-I protein fragment.

KW Apolipoprotein A-I; ApoA-I; cardiant; gene therapy; rabbit;
KW cardiovascular disorder.

OS Oryctolagus cuniculus.

PN WO2003035691-A1.

PD 01-MAY-2003.

PF 25-OCT-2002; 2002WO-CA001615.

PR 26-OCT-2001; 2001US-0335075P.

PA (XENO-) XENON GENETICS INC.

PA (UYBR-) UNIV BRITISH COLUMBIA.

PI Brownlie AJ, Dube M, Samuels M, Hayden MR;

DR WPI; 2003-449258/42.

PT New Apolipoprotein A-I mutation, useful for preparing a composition for
PT treating cardiovascular disorder.

PS Disclosure; Fig 4; 52pp; English.

CC The invention relates to a mutation of the apolipoprotein A-I (ApoA-I) comprising a leucine to proline variant in position 178. The mutant ApoA-I polynucleotide is useful for preparing a composition for treating cardiovascular disorder. Sequences ABR44037-44 represent ApoA-I fragments from various species where the leucine at position 178 is highly conserved except in pig (ABR44042)

XX
SQ Sequence 49 AA;

Query Match 33.5%; Score 186; DB 6; Length 49;
Best Local Similarity 77.6%; Pred. No. 1.5e-11;
Matches 38; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 26 DALRTHLAPYSDELQRQLARLEALKENGCGARLAELYHAKATEHLSTLSE 74
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 DTLRTHLAPYSDELQRQLARLEALKENGCGARLAELYHAKATEHLSTLSE 49

RESULT 6

ADC29656

ID ADC29656 standard; peptide; 33 AA.

XX ADC29656;

DT 18-DEC-2003 (first entry)

DE Antioxidant peptide #31.

KW antioxidant; cardiovascular disease; ischaemia; bone disease;
KW inflammatory related disease; human.

OS Synthetic.

OS Homo sapiens.

PN US2003087819-A1.

PD 08-MAY-2003.

PF 08-MAY-2002; 2002US-00142238.

PR 09-MAY-2001; 2001US-0289944P.

PA (BIEL/) BIELICKI J K.

XX Bielicki JK;

DR WPI; 2003-786946/74.

PT New cysteine-containing peptides, having antioxidant properties, useful
PT for diagnosing and treating cardiovascular disease, ischemia, bone
PT disease and other inflammatory related diseases.

PS Example 8; SEQ ID NO 31; 49pp; English.

CC The invention relates to an antioxidant peptide. The methods and
CC compositions of the present invention are useful for diagnosing and
CC treating cardiovascular disease, ischaemia, bone disease and other
CC inflammatory related diseases. The present sequence represents the amino
CC acid sequence of a peptide with antioxidant activity.

XX
SQ Sequence 33 AA;

Query Match 29.0%; Score 161; DB 7; Length 33;
Best Local Similarity 100.0%; Pred. No. 3.3e-09;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 78 PALEDLRQGLLPVLESFKVSFLSALREYTKLN 110
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 PALEDLRQGLLPVLESFKVSFLSALREYTKLN 33

RESULT 7

ABR44042
ID ABR44042 standard; peptide; 48 AA.

AC ABR44042;

DT 04-AUG-2003 (first entry)

DE Fig ApoA-I protein fragment.

KW Apolipoprotein A-I; ApoA-I; cardiant; gene therapy; pig;
KW cardiovascular disorder.

OS Sus sp.

PN WO2003035691-A1.

PD 01-MAY-2003.

PF 25-OCT-2002; 2002WO-CA001615.

PR 26-OCT-2001; 2001US-0335075P.

XX PA (XENO-) XENON GENETICS INC.
XX PI (UYBR-) UNIV BRITISH COLUMBIA.
XX PI Brownlie AJ, Dube M, Samuels M, Hayden MR;
XX DR WPI; 2003-449258/42.
XX PT New Apolipoprotein A-1 mutation, useful for preparing a composition for
XX PT treating cardiovascular disorder.
XX PS Disclosure; Fig 4; 52pp; English.
XX CC The invention relates to a mutation of the apolipoprotein A-I (ApoA-I)
XX CC comprising a leucine to proline variant in position 178. The mutant ApoA-
XX CC I polynucleotide is useful for preparing a composition for treating
XX CC cardiovascular disorder. Sequences ABR44037-44 represent ApoA-I fragments
XX CC from various species where the leucine at position 178 is highly
XX CC conserved except in pig (ABR44042)
XX SQ Sequence 48 AA;

Query Match 28.3%; Score 157.5; DB 6; Length 48;
Best Local Similarity 69.4%; Pred. No. 1.2e-08;
Matches 34; Conservative 4; Mismatches 10; Indels 1; Gaps 1;

QY 26 DALRTHLAPYSDELQRLAARLEALKENGCGARLAHYHAKATEHLSTLSE 74
Db 1 EALRQQLAPYSDDLRQRMARFALKEGGGS-LAEYQAKAQEQLKALGE 48

RESULT 8
ADC29655
ID ADC29655 standard; peptide; 33 AA.
XX AC ADC29655;
XX DT 18-DEC-2003 (first entry)
XX DE Antioxidant peptide #30.
XX KW antioxidant; cardiovascular disease; ischaemia; bone disease;
XX KW inflammatory related disease; human.
XX OS Synthetic.
XX OS Homo sapiens.
XX PN US2003087819-A1.
XX PD 08-MAY-2003.
XX PF 08-MAY-2002; 2002US-00142238.
XX PR 09-MAY-2001; 2001US-0289944P.
XX PA (BIEL/) BIELICKI J K.
XX PI Bielicki JK;
XX DR WPI; 2003-786946/74.
XX PT New cysteine-containing peptides, having antioxidant properties, useful
XX PT for diagnosing and treating cardiovascular disease, ischemia, bone
XX PT disease and other inflammatory related diseases.
XX PS Example 8; SEQ ID NO 30; 49pp; English.
XX CC The invention relates to an antioxidant peptide. The methods and
XX CC compositions of the present invention are useful for diagnosing and
XX CC treating cardiovascular disease, ischaemia, bone disease and other
XX CC inflammatory related diseases. The present sequence represents the amino
XX CC acid sequence of a peptide with antioxidant activity.

SQ Sequence 33 AA;

Query Match 27.5%; Score 153; DB 7; Length 33;
Best Local Similarity 97.0%; Pred. No. 2.2e-08;
Matches 32; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 78 PALEDLROGLLPVLESFKVSLSALEEYTKLN 110
Db 1 PALEDLROGLLPVLESFCVSLSALEEYTKLN 33

RESULT 9
ABR44044
ID ABR44044 standard; peptide; 49 AA.
XX AC ABR44044;
XX DT 04-AUG-2003 (first entry)
XX DE Chicken ApoA-I protein fragment.
XX KW Apolipoprotein A-I; ApoA-I; cardiant; gene therapy; chicken;
XX KW cardiovascular disorder.
XX OS Gallus sp.
XX PN WO2003035691-A1.
XX PD 01-MAY-2003.
XX PF 25-OCT-2002; 2002WO-CA001615.
XX PR 26-OCT-2001; 2001US-0335075P.
XX PA (XENO-) XENON GENETICS INC.
XX PA (UYBR-) UNIV BRITISH COLUMBIA.
XX PI Brownlie AJ, Dube M, Samuels M, Hayden MR;
XX DR WPI; 2003-449258/42.
XX PT New Apolipoprotein A-1 mutation, useful for preparing a composition for
XX PT treating cardiovascular disorder.
XX PS Disclosure; Fig 4; 52pp; English.
XX CC The invention relates to a mutation of the apolipoprotein A-I (ApoA-I)
XX CC comprising a leucine to proline variant in position 178. The mutant ApoA-
XX CC I polynucleotide is useful for preparing a composition for treating
XX CC cardiovascular disorder. Sequences ABR44037-44 represent ApoA-I fragments
XX CC from various species where the leucine at position 178 is highly
XX CC conserved except in pig (ABR44042)
XX SQ Sequence 49 AA;

Query Match 23.2%; Score 129; DB 6; Length 49;
Best Local Similarity 53.1%; Pred. No. 1e-05;
Matches 26; Conservative 9; Mismatches 14; Indels 0; Gaps 0;

QY 26 DALRTHLAPYSDELQRLAARLEALKENGCGARLAHYHAKATEHLSTLSE 74
Db 1 ESELKRLAPYSDELQRLKLSQKLEIREKGIPOASEYQAKVMEQLNLRE 49

RESULT 10
ABR44043
ID ABR44043 standard; peptide; 47 AA.
XX AC ABR44043;
XX DT 04-AUG-2003 (first entry)
XX DE Rat ApoA-I protein fragment.

Result No.	Score	Query			ID	Description
		Match	Length	DB		
1	161	29.0	33	14	US-10-142-238A-31	Sequence 31, Appl
2	153	27.5	33	14	US-10-142-238A-30	Sequence 30, Appl
3	117	21.0	24	15	US-10-465-789A-58	Sequence 58, Appl
4	113	20.3	22	15	US-10-465-789A-54	Sequence 54, Appl
5	110	19.8	22	15	US-10-465-789A-55	Sequence 55, Appl
6	108	19.4	22	15	US-10-465-789A-56	Sequence 56, Appl
7	94	16.9	18	14	US-10-142-238A-10	Sequence 10, Appl
8	90	16.2	18	14	US-10-142-238A-13	Sequence 13, Appl
9	87	15.6	18	14	US-10-142-238A-17	Sequence 17, Appl
10	86	15.5	18	14	US-10-142-238A-9	Sequence 9, Appl
11	86	15.5	18	14	US-10-142-238A-11	Sequence 11, Appl
12	86	15.5	18	14	US-10-142-238A-14	Sequence 14, Appl
13	85	15.3	18	14	US-10-142-238A-12	Sequence 12, Appl

US-10-142-238A-31

; Sequence 30, Application US/10142238A

```
; Publication No. US2003008781A1
; GENERAL INFORMATION:
; APPLICANT: Bielski, John K.
; TITLE OF INVENTION: CYSTEINE-CONTAINING PEPTIDES HAVING OXIDANT PROPERTIES
; FILE REFERENCE: IB-1705
; CURRENT APPLICATION NUMBER: US/10/142,238A
; CURRENT FILING DATE: 2002-08-19
; PRIOR APPLICATION NUMBER: US 60/289,944
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 30
; LENGTH: 33
; TYPE: PRT
; ORGANISM: ARTIFICIAL SEQUENCE
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(33)
; OTHER INFORMATION: HUMAN GENETIC ORIGIN
US-10-142-238A-30

Query Match      27.5%; Score 153; DB 14; Length 33;
Best Local Similarity 97.0%; Pred. No. 8.3e-09;
Matches 32; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 78 PALEDLROGLLPVLESFKVSLSALEYTKLN 110
   |||||
Db 1 PALEDLROGLLPVLESFKVSLSALEYTKLN 33

RESULT 3
US-10-465-789A-58
; Sequence 58, Application US/10465789A
; Publication No. US20040053384A1
; GENERAL INFORMATION:
; APPLICANT: Sligar, Stephen G
; APPLICANT: Bayburt, Timothy H
; APPLICANT: Schuler, Mary A
; APPLICANT: Civjan, Natanya R
; APPLICANT: Yelena V. Grinkova
; APPLICANT: Ilia G. Denisov
; TITLE OF INVENTION: Membrane Scaffold Proteins
; FILE REFERENCE: 87-00A
; CURRENT APPLICATION NUMBER: US/10/465,789A
; CURRENT FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: 09/990,087
; PRIOR FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 60/252,233
; PRIOR FILING DATE: 2000-11-20
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 58
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: artificial sequence of Helix 10
US-10-465-789A-58

Query Match      21.0%; Score 117; DB 15; Length 24;
Best Local Similarity 100.0%; Pred. No. 3.2e-05;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 89 PVLESFKVSLSALEYTKLNTQ 112
   |||||
Db 1 PVLESFKVSLSALEYTKLNTQ 24

RESULT 4
US-10-465-789A-54
; Sequence 54, Application US/10465789A
; Publication No. US20040053384A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Sligar, Stephen G
; APPLICANT: Bayburt, Timothy H
; APPLICANT: Schuler, Mary A
; APPLICANT: Civjan, Natanya R
; APPLICANT: Yelena V. Grinkova
; APPLICANT: Ilia G. Denisov
; TITLE OF INVENTION: Membrane Scaffold Proteins
; FILE REFERENCE: 87-00A
; CURRENT APPLICATION NUMBER: US/10/465,789A
; CURRENT FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: 09/990,087
; PRIOR FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 60/252,233
; PRIOR FILING DATE: 2000-11-20
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 54
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: artificial sequence of Helix 6
US-10-465-789A-54

Query Match      20.3%; Score 113; DB 15; Length 22;
Best Local Similarity 100.0%; Pred. No. 7.4e-05;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 PLGEEMDRARAHVDALRTHLA 33
   |||||
Db 1 PLGEEMDRARAHVDALRTHLA 22

RESULT 5
US-10-465-789A-55
; Sequence 55, Application US/10465789A
; Publication No. US20040053384A1
; GENERAL INFORMATION:
; APPLICANT: Sligar, Stephen G
; APPLICANT: Bayburt, Timothy H
; APPLICANT: Schuler, Mary A
; APPLICANT: Civjan, Natanya R
; APPLICANT: Yelena V. Grinkova
; APPLICANT: Ilia G. Denisov
; TITLE OF INVENTION: Membrane Scaffold Proteins
; FILE REFERENCE: 87-00A
; CURRENT APPLICATION NUMBER: US/10/465,789A
; CURRENT FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: 09/990,087
; PRIOR FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 60/252,233
; PRIOR FILING DATE: 2000-11-20
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 55
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: artificial sequence of Helix 7
US-10-465-789A-55

Query Match      19.8%; Score 110; DB 15; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.00015;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 PYSDELRLRLAARLEALKENG 55
   |||||
Db 1 PYSDELRLRLAARLEALKENG 22

RESULT 6
US-10-465-789A-56
```

```
; Sequence 56, Application US/10465789A
; Publication No. US20040053384A1
; GENERAL INFORMATION:
; APPLICANT: Sliagar, Stephen G
; APPLICANT: Bayburt, Timothy H
; APPLICANT: Schuler, Mary A
; APPLICANT: Civjan, Natanya R
; APPLICANT: Yelena V. Grinkova
; APPLICANT: Iliia G. Denisov
; TITLE OF INVENTION: Membrane Scaffold Proteins
; FILE REFERENCE: 87-00A
; CURRENT APPLICATION NUMBER: US/10/465,789A
; CURRENT FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: 09/990,087
; PRIOR FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 60/252,233
; PRIOR FILING DATE: 2000-11-20
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 56
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; LOCATION: (1)..(18)
; OTHER INFORMATION: artificial sequence of Helix 8
US-10-465-789A-56

Query Match      19.4%; Score 108; DB 15; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.0025;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 ARLAEYHAKATEHLSTLSEKAK 77
      |||||
Db 1 ARLAEYHAKATEHLSTLSEKAK 22

RESULT 7
US-10-142-238A-10
; Sequence 10, Application US/10142238A
; Publication No. US20030087819A1
; GENERAL INFORMATION:
; APPLICANT: Bielicki, John K.
; TITLE OF INVENTION: CYSTEINE-CONTAINING PEPTIDES HAVING OXIDANT PROPERTIES
; FILE REFERENCE: IB-1705
; CURRENT APPLICATION NUMBER: US/10/142,238A
; CURRENT FILING DATE: 2002-08-19
; PRIOR APPLICATION NUMBER: US 60/289,944
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 18
; TYPE: PRT
; ORGANISM: ARTIFICIAL SEQUENCE
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(18)
; OTHER INFORMATION: HUMAN GENETIC ORIGIN
US-10-142-238A-10

Query Match      16.9%; Score 94; DB 14; Length 18;
Best Local Similarity 100.0%; Pred. No. 0.0056;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 GEEMDRARAHVDALRTH 31
      |||||
Db 1 GEEMDRARAHVDALRTH 18

RESULT 8
US-10-142-238A-13
; Sequence 13, Application US/10142238A
; Publication No. US20030087819A1
; GENERAL INFORMATION:
; APPLICANT: Bielicki, John K.
; TITLE OF INVENTION: CYSTEINE-CONTAINING PEPTIDES HAVING OXIDANT PROPERTIES
; FILE REFERENCE: IB-1705
; CURRENT APPLICATION NUMBER: US/10/142,238A
; CURRENT FILING DATE: 2002-08-19
; PRIOR APPLICATION NUMBER: US 60/289,944
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 18
; TYPE: PRT
; ORGANISM: ARTIFICIAL SEQUENCE
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(18)
; OTHER INFORMATION: HUMAN GENETIC ORIGIN
US-10-142-238A-13

Query Match      16.2%; Score 90; DB 14; Length 18;
Best Local Similarity 94.4%; Pred. No. 0.015;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 14 GEEMDRARAHVDALRTH 31
      |||||
Db 1 GEEMDRARAHVDALRTH 18

RESULT 9
US-10-142-238A-17
; Sequence 17, Application US/10142238A
; Publication No. US20030087819A1
; GENERAL INFORMATION:
; APPLICANT: Bielicki, John K.
; TITLE OF INVENTION: CYSTEINE-CONTAINING PEPTIDES HAVING OXIDANT PROPERTIES
; FILE REFERENCE: IB-1705
; CURRENT APPLICATION NUMBER: US/10/142,238A
; CURRENT FILING DATE: 2002-08-19
; PRIOR APPLICATION NUMBER: US 60/289,944
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 18
; TYPE: PRT
; ORGANISM: ARTIFICIAL SEQUENCE
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(18)
; OTHER INFORMATION: HUMAN GENETIC ORIGIN
US-10-142-238A-17

Query Match      15.6%; Score 87; DB 14; Length 18;
Best Local Similarity 100.0%; Pred. No. 0.03;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 89 PVLESFKVSFLSALREYT 106
      |||||
Db 1 PVLESFKVSFLSALREYT 18

RESULT 10
US-10-142-238A-9
; Sequence 9, Application US/10142238A
; Publication No. US20030087819A1
; GENERAL INFORMATION:
; APPLICANT: Bielicki, John K.
; TITLE OF INVENTION: CYSTEINE-CONTAINING PEPTIDES HAVING OXIDANT PROPERTIES
; FILE REFERENCE: IB-1705
; CURRENT APPLICATION NUMBER: US/10/142,238A
; CURRENT FILING DATE: 2002-08-19
; PRIOR APPLICATION NUMBER: US 60/289,944
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 18
; TYPE: PRT
; ORGANISM: ARTIFICIAL SEQUENCE
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(18)
; OTHER INFORMATION: HUMAN GENETIC ORIGIN
US-10-142-238A-9

Query Match      15.6%; Score 87; DB 14; Length 18;
Best Local Similarity 100.0%; Pred. No. 0.03;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 89 PVLESFKVSFLSALREYT 106
      |||||
Db 1 PVLESFKVSFLSALREYT 18

RESULT 10
US-10-142-238A-9
; Sequence 9, Application US/10142238A
; Publication No. US20030087819A1
; GENERAL INFORMATION:
; APPLICANT: Bielicki, John K.
; TITLE OF INVENTION: CYSTEINE-CONTAINING PEPTIDES HAVING OXIDANT PROPERTIES
; FILE REFERENCE: IB-1705
; CURRENT APPLICATION NUMBER: US/10/142,238A
; CURRENT FILING DATE: 2002-08-19
; PRIOR APPLICATION NUMBER: US 60/289,944
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 18
; TYPE: PRT
; ORGANISM: ARTIFICIAL SEQUENCE
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(18)
; OTHER INFORMATION: HUMAN GENETIC ORIGIN
US-10-142-238A-9
```

; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 18
; TYPE: PRT
; ORGANISM: ARTIFICIAL SEQUENCE
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(18)
; OTHER INFORMATION: HUMAN GENETIC ORIGIN
US-10-142-238A-9

Query Match 15.5%; Score 86; DB 14; Length 18;
Best Local Similarity 94.4%; Pred. No. 0.038;
Matches 17; Conservative 0; Mismatches 0; Gaps 0;

QY 14 GEEMDRARAHVDALRTH 31
||||| ||||| ||||| |||||
Db 1 GEEMDCARAHVDALRTH 18

RESULT 11
US-10-142-238A-11
; Sequence 11, Application US/10142238A
; Publication No. US20030087819A1
; GENERAL INFORMATION:
; APPLICANT: Bielicki, John K.
; TITLE OF INVENTION: CYSTEINE-CONTAINING PEPTIDES HAVING OXIDANT PROPERTIES
; FILE REFERENCE: IB-1705
; CURRENT APPLICATION NUMBER: US/10/142,238A
; CURRENT FILING DATE: 2002-08-19
; PRIOR APPLICATION NUMBER: US 60/289,944
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 18
; TYPE: PRT
; ORGANISM: ARTIFICIAL SEQUENCE
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(18)
; OTHER INFORMATION: HUMAN GENETIC ORIGIN
US-10-142-238A-11

Query Match 15.5%; Score 86; DB 14; Length 18;
Best Local Similarity 94.4%; Pred. No. 0.038;
Matches 17; Conservative 0; Mismatches 0; Gaps 0;

QY 14 GEEMDRARAHVDALRTH 31
||||| ||||| ||||| |||||
Db 1 GEEMCDARAHVDALRTH 18

RESULT 12
US-10-142-238A-14
; Sequence 14, Application US/10142238A
; Publication No. US20030087819A1
; GENERAL INFORMATION:
; APPLICANT: Bielicki, John K.
; TITLE OF INVENTION: CYSTEINE-CONTAINING PEPTIDES HAVING OXIDANT PROPERTIES
; FILE REFERENCE: IB-1705
; CURRENT APPLICATION NUMBER: US/10/142,238A
; CURRENT FILING DATE: 2002-08-19
; PRIOR APPLICATION NUMBER: US 60/289,944
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 18
; TYPE: PRT
; ORGANISM: ARTIFICIAL SEQUENCE

; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(18)
; OTHER INFORMATION: HUMAN GENETIC ORIGIN
US-10-142-238A-14

Query Match 15.5%; Score 86; DB 14; Length 18;
Best Local Similarity 94.4%; Pred. No. 0.038;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 14 GEEMDRARAHVDALRTH 31
||||| ||||| ||||| |||||
Db 1 GEEMDRACAHVDALRTH 18

RESULT 13
US-10-142-238A-12
; Sequence 12, Application US/10142238A
; Publication No. US20030087819A1
; GENERAL INFORMATION:
; APPLICANT: Bielicki, John K.
; TITLE OF INVENTION: CYSTEINE-CONTAINING PEPTIDES HAVING OXIDANT PROPERTIES
; FILE REFERENCE: IB-1705
; CURRENT APPLICATION NUMBER: US/10/142,238A
; CURRENT FILING DATE: 2002-08-19
; PRIOR APPLICATION NUMBER: US 60/289,944
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 18
; TYPE: PRT
; ORGANISM: ARTIFICIAL SEQUENCE
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(18)
; OTHER INFORMATION: HUMAN GENETIC ORIGIN
US-10-142-238A-12

Query Match 15.3%; Score 85; DB 14; Length 18;
Best Local Similarity 94.4%; Pred. No. 0.048;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 14 GEEMDRARAHVDALRTH 31
||||| ||||| ||||| |||||
Db 1 GEEMCRARAHVDALRTH 18

RESULT 14
US-10-142-238A-2
; Sequence 2, Application US/10142238A
; Publication No. US20030087819A1
; GENERAL INFORMATION:
; APPLICANT: Bielicki, John K.
; TITLE OF INVENTION: CYSTEINE-CONTAINING PEPTIDES HAVING OXIDANT PROPERTIES
; FILE REFERENCE: IB-1705
; CURRENT APPLICATION NUMBER: US/10/142,238A
; CURRENT FILING DATE: 2002-08-19
; PRIOR APPLICATION NUMBER: US 60/289,944
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 18
; TYPE: PRT
; ORGANISM: ARTIFICIAL SEQUENCE
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(18)
; OTHER INFORMATION: HUMAN GENETIC ORIGIN
US-10-142-238A-2

Query Match 15.1%; Score 84; DB 14; Length 18;

Best Local Similarity 100.0%; Pred. No. 0.062; Mismatches 0; Indels 0; Gaps 0;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 SDELRLRLAARLEALKEN 53
| | | | | | | | | | | | | | | | | |
Db 1 SDELRLRLAARLEALKEN 18

RESULT 15
US-10-142-238A-15
; Sequence 15, Application US/10142238A
; Publication No. US20030087819A1
; GENERAL INFORMATION:
; APPLICANT: Bielsicki, John K.
; TITLE OF INVENTION: CYSTEINE-CONTAINING PEPTIDES HAVING OXIDANT PROPERTIES
; FILE REFERENCE: IB-1705
; CURRENT APPLICATION NUMBER: US/10/142,238A
; CURRENT FILING DATE: 2002-08-19
; PRIOR APPLICATION NUMBER: US 60/289,944
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 18
; TYPE: PRT
; ORGANISM: ARTIFICIAL SEQUENCE
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(18)
; OTHER INFORMATION: HUMAN GENETIC ORIGIN
US-10-142-238A-15

Query Match 14.9%; Score 83; DB 14; Length 18;
Best Local Similarity 94.4%; Pred. No. 0.078; Mismatches 1; Indels 0; Gaps 0;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 14 GEEMDRARAHVDALRTH 31
| | | | | | | | | | | | | | | | | |
Db 1 GEEMDRARACVDALRTH 18

Search completed: December 21, 2004, 11:46:57
Job time : 60 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2004, 11:46:06 ; Search time 38 Seconds
(without alignments)
195.464 Million cell updates/sec

Title: US-09-803-918a-2_COPY_156_267

Perfect score: 556
Sequence: 1 QXHEHQEKLSPGEMRDR.....SPKVSFLSALEYTKKLNTQ 112

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 320001

Minimum DB seq length: 0
Maximum DB seq length: 112

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	86	15.5	64	US-08-292-870-1	Sequence 1, Appli
2	80.5	14.5	110	US-07-849-389-7	Sequence 7, Appli
3	68.5	12.3	100	US-09-147-875A-2	Sequence 2, Appli
4	67.5	12.1	70	5182364-4	Patent No. 5182364
5	67.5	12.1	101	US-08-710-749-1	Sequence 1, Appli
6	66.5	12.0	100	US-09-147-875A-3	Sequence 3, Appli
7	66.5	12.0	101	US-08-710-749-2	Sequence 2, Appli
8	66.5	12.0	104	US-08-777-708C-5	Sequence 5, Appli
9	65	11.7	69	5168045-4	Patent No. 5168045
10	64.5	11.6	70	5177189-3	Patent No. 5177189
11	63.5	11.4	100	US-09-147-875A-4	Sequence 4, Appli
12	62.5	11.2	101	US-08-710-749-3	Sequence 3, Appli
13	62.5	11.2	109	US-08-893-042-4	Sequence 4, Appli
14	60.5	10.9	100	US-09-147-875A-5	Sequence 5, Appli
15	59.5	10.7	77	US-08-182-175A-57	Sequence 57, Appl
16	59.5	10.7	77	US-08-474-633A-75	Sequence 75, Appl
17	59.5	10.7	77	US-08-823-771-75	Sequence 75, Appl
18	59.5	10.7	77	PCT-US92-06412-57	Sequence 57, Appl
19	59.5	10.7	101	US-08-710-749-4	Sequence 4, Appli
20	59	10.6	40	US-09-214-742-2	Sequence 2, Appli
21	59	10.6	40	US-09-255-331-2	Sequence 2, Appli
22	59	10.6	41	US-08-584-043A-76	Sequence 76, Appl
23	59	10.6	70	US-09-489-039A-8458	Sequence 8458, Ap
24	59	10.6	94	US-09-340-820A-67	Sequence 67, Appl
25	59	10.6	99	US-09-147-875A-13	Sequence 13, Appl
26	59	10.6	108	US-08-710-749-25	Sequence 25, Appl
27	59	10.6	108	US-09-147-875A-26	Sequence 26, Appl

28	58.5	10.5	107	1	US-08-182-175A-105	Sequence 105, App
29	58.5	10.5	107	1	US-08-474-633A-92	Sequence 92, Appl
30	58.5	10.5	107	4	US-08-823-771-92	Sequence 92, Appl
31	58.5	10.5	107	5	PCT-US92-06412-105	Sequence 105, App
32	58	10.4	101	2	US-08-710-749-5	Sequence 5, Appli
33	57.5	10.3	100	4	US-09-147-875A-12	Sequence 12, Appl
34	57.5	10.3	103	4	US-09-513-999C-8123	Sequence 8123, Ap
35	57.5	10.3	106	4	US-09-252-991A-24884	Sequence 24884, A
36	57	10.3	96	4	US-09-583-110-3341	Sequence 3341, App
37	57	10.3	105	4	US-09-071-035-292	Sequence 292, App
38	57	10.3	110	4	US-09-270-767-58246	Sequence 58246, A
39	56.5	10.2	98	4	US-09-147-875A-1	Sequence 1, Appli
40	56.5	10.2	100	4	US-09-147-875A-6	Sequence 6, Appli
41	56.5	10.2	103	4	US-09-270-767-34122	Sequence 34122, A
42	56.5	10.2	103	4	US-09-270-767-49339	Sequence 49339, A
43	56.5	10.2	108	4	US-09-270-767-35185	Sequence 35185, A
44	56.5	10.2	108	4	US-09-270-767-50402	Sequence 50402, A
45	56	10.1	103	4	US-09-252-991A-29532	Sequence 29532, A

ALIGNMENTS

RESULT 1
US-08-292-870-1
; Sequence 1, Application US/08292870
; Patent No. 5814467
; GENERAL INFORMATION:
; APPLICANT: Curtiss, Linda K
; APPLICANT: Banks, Carole L
; APPLICANT: Bonnet, David J
; APPLICANT: Smith, Richard S
; TITLE OF INVENTION: APO AI POLYPEPTIDES, DIAGNOSTIC METHODS
; TITLE OF INVENTION: AND SYSTEMS FOR QUANTIFYING APO AI, AND THERAPEUTIC
; TITLE OF INVENTION: METHODS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSES: The Scripps Research Institute, Office of
; ADDRESSES: Patent Counsel
; STREET: 10666 No. 5814467th Torrey Pines Road., TPC 8
; CITY: La Jolla
; STATE: California
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/292,870
; FILING DATE: 17-AUG-1994
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/534,761
; FILING DATE: 07-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/711,333
; FILING DATE: 06-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US 91/04038
; FILING DATE: 07-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: TSRI 210.1 D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 64 amino acids
; TYPE: amino acid

```
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-08-292-870-1

Query Match      15.5%; Score 86; DB 2; Length 64;
Best Local Similarity 100.0%; Pred. No. 0.025;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QKLHELQEKLSPLGEEM 17
Db 48 QKLHELQEKLSPLGEEM 64

RESULT 2
US-07-849-389-7
; Sequence 7, Application US/07849389
; Patent No. 5525493
; GENERAL INFORMATION:
; APPLICANT: HORNES, Erik
; APPLICANT: UHLEN, Mathias
; TITLE OF INVENTION: CLONING METHOD AND KIT
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/849,389
; FILING DATE: 19920519
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16787/168/DFBC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-9300
; TELEFAX: (703) 683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 110 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-849-389-7

Query Match      14.5%; Score 80.5; DB 1; Length 110;
Best Local Similarity 30.4%; Pred. No. 0.21;
Matches 28; Conservative 13; Mismatches 29; Indels 29; Gaps 3;

QY 5 ELQEKLSPLGEEMRDRAR 42
Db 9 ELEEQLTVAEETRLRSKELQAAEPLGADMEDVGRVLVQYRGEVQAMLGQSTELRYR 68

QY 43 LAARLEALK-----ENGARLAIEYHAKATE 67
Db 69 LASHLRKLRKRLRDADDLQKRLAVYQAGARE 100

RESULT 3
US-09-147-875A-2
; Sequence 2, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
```

```
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO: 2
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-147-875A-2

Query Match      12.3%; Score 68.5; DB 4; Length 100;
Best Local Similarity 26.7%; Pred. No. 3.7;
Matches 23; Conservative 16; Mismatches 34; Indels 13; Gaps 3;

QY 2 KLHELQEKLSPLGEEMRDRARAHVDALRTHLAPYSDELQRQLAARLEALKENGARLAIEY 61
Db 24 ELDAKQAKLSKL-EELSDK----IDELDAEIA-----KLEKVEDFKNSNGEQAEQY 70

QY 62 HAKATEHLSTLSEKAKPALEDLRQGL 87
Db 71 RAAAEEDLAAKQAELEKTEADLKAV 96

RESULT 4
5182364-4
; Patent No. 5182364
; APPLICANT: DYER, CHERYL A.; CURTISS, LINDA K.; SMITH, RICHARD
; TITLE OF INVENTION: POLYPEPTIDE ANALOGS OF APOLIPOPROTEIN E
; NUMBER OF SEQUENCES: 14
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,158
; FILING DATE: 26-FEB-1990
; SEQ ID NO: 4
; LENGTH: 70
5182364-4

Query Match      12.1%; Score 67.5; DB 6; Length 70;
Best Local Similarity 27.4%; Pred. No. 3;
Matches 20; Conservative 13; Mismatches 31; Indels 9; Gaps 1;

QY 15 EEMDRARAHVDALRTHLAPYSDELQRQLAARLEALKENGARLAIEYHAKATEHLSTLSE 74
Db 3 EELRVRLASHLRKLRKRLRDADDLQKRLAVYQST-----BELRVRLASHLRKLRK 53

QY 75 KAKPALEDLRQGL 87
Db 54 RLRLRDADDLQKRL 66

RESULT 5
US-08-710-749-1
; Sequence 1, Application US/08710749
; Patent No. 5955089
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Becker, Robert
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE
; PROTEINS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/710,749
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2074
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 101 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
; US-08-710-749-1

Query Match 12.1%; Score 67.5; DB 2; Length 101;
Best Local Similarity 27.2%; Pred. No. 4.9;
Matches 22; Conservative 15; Mismatches 31; Indels 13; Gaps 3;

QY 7 QEKLSPLGEMDRARAHVDALRTHLAPYSDELQRLAARLEALKENGARLAEYHAKAT 66
Db 30 QAKLSKL-BELSDK----IDELDAEIA-----KLEKNVEDFKNSNGEQAYRAAE 76
QY 67 EHLSTLSEKAKPALEDLRQGL 87
Db 77 EDLAQAQAELEKTEADLKAV 97

RESULT 6
US-09-147-875A-3
; Sequence 3, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; US-09-147-875A-3

Query Match 12.0%; Score 66.5; DB 4; Length 100;
Best Local Similarity 26.7%; Pred. No. 6.2;
Matches 23; Conservative 15; Mismatches 35; Indels 13; Gaps 3;

QY 2 KLHELEKLSPLGEMDRARAHVDALRTHLAPYSDELQRLAARLEALKENGARLAEY 61
Db 24 ELDAKQAKLSKL-BELSDK----IDELDAEIA-----KLEKNVEDFKNSNGEAEQY 70
QY 62 HAKATEHLSTLSEKAKPALEDLRQGL 87
Db 71 RAAAGEDLAQAQAELEKTEADLKAV 96

RESULT 7
US-08-710-749-2
; Sequence 2, Application US/08710749
; Patent No. 5955089
; GENERAL INFORMATION:
; APPLICANT: Brilles, David B.
; APPLICANT: Hollingshead, Susan
```

```

; APPLICANT: Becker, Robert
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE
; PROTEINS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/710,749
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2074
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 101 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
; US-08-710-749-2

Query Match 12.0%; Score 66.5; DB 2; Length 101;
Best Local Similarity 27.2%; Pred. No. 6.3;
Matches 22; Conservative 15; Mismatches 31; Indels 13; Gaps 3;

QY 7 QEKLSPLGEMDRARAHVDALRTHLAPYSDELQRLAARLEALKENGARLAEYHAKAT 66
Db 30 QAKLSKL-BELSDK----IDELDAEIA-----KLEKNVEDFKNSNGEQAYRAAAG 76
QY 67 EHLSTLSEKAKPALEDLRQGL 87
Db 77 EDLAQAQAELEKTEADLKAV 97

RESULT 8
US-08-777-708C-5
; Sequence 5, Application US/08777708C
; Patent No. 6159934
; GENERAL INFORMATION:
; APPLICANT: Pescovitz, Ora H.
; TITLE OF INVENTION: USE OF GHRH-RP TO STIMULATE STEM CELL FACTOR PRODUCTION
; FILE REFERENCE: 7037-175/IU-62
; CURRENT APPLICATION NUMBER: US/08/777,708C
; CURRENT FILING DATE: 1996-12-20
; PRIOR APPLICATION NUMBER: US 60/008,933
; PRIOR FILING DATE: 1995-12-20
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; US-08-777-708C-5

Query Match 12.0%; Score 66.5; DB 3; Length 104;
Best Local Similarity 32.9%; Pred. No. 6.5;
Matches 23; Conservative 10; Mismatches 32; Indels 5; Gaps 3;
```

Qy	20	RARAHVDALATHAPYSDELRQRLAARL--EALKENGGAARLARYAHAKATEHJLSTL--SEKA	76
Db	27	RVRHADAIPT--SSYRRLGQLYARKLLHEIMNRQGGERNQEQSRFRNRHLDRVMAEDK	84
Qy	77	KPALEDLRQG	86
Db	85	OMALESILQG	94

```

RESULT 9
5168045-4
; Patent No. 5168045
; APPLICANT: DYER, CHERYL A.; CURTISS, LINDA K.; SMITH,
; RICHARD
; TITLE OF INVENTION: DIAGNOSTIC SYSTEMS AND METHODS USING
; POLYPEPTIDE ANALOGS OF APOLIPOPROTEIN E
; NUMBER OF SEQUENCES: 8
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/540,363
; FILING DATE: 18-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 485,158
; FILING DATE: 26-FEB-1990
; APPLICATION NUMBER: 595,732
; FILING DATE: 18-AUG-1989
; SEQ ID NO:4
; LENGTH: 69
5168045-4

```

```

Query Match      11.7%; Score 65; DB 6; Length 69;
Best Local Similarity 28.8%; Pred. No. 5.5;
Matches 21; Conservative 13; Mismatches 29; Indels 10; Gaps 2;

Qy 15 EEMDRARAHVDALRTHLAPYSDLRQLRKLRLDADDDLRKLVAYQST-----EELRVRLASHLRKLK 74
Db 3 BELRVRLASHLRKLRLDADDDLRKLVAYQST-----EELRVRLASHLRKLK 53

Qy 75 KAKPALEDLRQGL 87
Db 54 LLRDA-DLQKRL 65

```

```

RESULT 10
5177189-3
;Patent No. 5177189
; APPLICANT: DYER, CHERYL A.;CURTISS, LINDA K.;SMITH, RICHARD
; TITLE OF INVENTION: POLYPEPTIDE ANALOGS OF APOLIPROTEIN E
; NUMBER OF SEQUENCES: 11
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/395,732
; FILING DATE: 18-AUG-1989
; SEQ ID NO:3
; LENGTH: 70
5177189-3

```

```
Query Match      11.6%; Score 64.5; DB 6; Length 70;
Best Local Similarity 27.4%; Pred.No. 6.4;
Matches 20; Conservative 12; Mismatches 32; Indels 9; Gaps 1;
```

Qy	15	EEMDRARAHVDALRTHLAPYSDDELRLQRLAARLEALKENGGAFLAEYHAKATEHLSTLSE	74
	:	:	:
Db	3	EELVRLASHLRKURKLRRDADDLOKRELAYVYST-----EELRVTLASHLRKLRK	53
		:	:
Qy	75	KAKPALEDLROGL	87
		:	:
Db	54	RLLRDADDLOKRL	66
		:	:

RESULT 11
US-09-147-875A-4
; Sequence 4, Application US/09147875A
; Patent No. 663816
; GENERAL INFORMATION:

```

; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; US-09-147-875A-4

```

	Query Match	11.4%	Score 63.5	DB 4	Length 100
Best Local Similarity	27.9%	Fred.No.13			
Matches	24	Conservative	34	Indels 13	Gaps 3
Qy	2	KLHELOEKLSPGELMRDRAHVDAURLTHLAPVSDLRQLAARLFALEKXNGGARLAEY	61		
Db	24	ELDKAKLKL-BELSDK----	IDELDABIA-----	KLEKDVDFKNSDGEAQGY	70
Qy	62	HAKATEHLSTLSKAKPALEDRLQGL	87		
Db	71	LAAGGEDLIAKKAELEKAEADLKAV	96		

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Query Match	11.2%;	Score 62.5;	DB 2;	Length 101;
Best Local Similarity	28.4%;	Pred. No. 17;		
Matches 23;	Conservative 14;	Mismatches 31;	Indels 13;	Gaps 3;

5

QY 7 QEKLSPGEMRDRARAHVDALRTHLAPYSDELQRLAARLEALKENGARLAAYHAKAT 66
Db 30 QAKLSKL-EELSJK-----IDELDAEIA-----KLEKDVDFKNSDGEQAGYLAAAG 76
QY 67 EHLSTLSEKAKPALEDLRQL 87
Db 77 EDLIAKAELEKAEADLUKAV 97

RESULT 13
US-08-893-042-4
; Sequence 4, Application US/08893042
; Patent No. 5906923
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; TITLE OF INVENTION: NOVEL ATPASE INHIBITOR
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/893,042
; FILING DATE: Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0134 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 109 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 162713
US-08-893-042-4

Query Match 11.2%; Score 62.5; DB 2; Length 109;
Best Local Similarity 27.5%; Pred. No. 19;
Matches 22; Conservative 13; Mismatches 36; Indels 9; Gaps 2;

QY 14 GEWRDRARAHVDALR-----THLAPYSDELQRLAARLEALKENGARLAAYHAKA 65
Db 26 GSSESGDNVRSAGVFRAGKQREAEERYFRAPAKQQLAALKKHENEIS-HHAK 84

QY 66 TEHLSTLSEKAKPALEDLRQ 85
Db 85 IERLQKEIERHKOSIKLKQ 104

RESULT 14
US-09-147-875A-5
; Sequence 5, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.

; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; CURRENT FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-147-875A-5

Query Match 10.9%; Score 60.5; DB 4; Length 100;
Best Local Similarity 26.7%; Pred. No. 28;
Matches 23; Conservative 15; Mismatches 35; Indels 13; Gaps 3;

QY 2 KLHELQEKLSPLGEMRDRARAHVDALRTHLAPYSDELQRLAARLEALKENGARLAAY 61
Db 24 ELDKQAKLSKL-EELSDK----IDELDAEIA-----KLEKDVDFKNSDGEQAGQY 70

QY 62 HAKATEHLSTLSEKAKPALEDLRQL 87
Db 71 LAAAEEDLIAKAELEQTEADLUKAV 96

RESULT 15
US-08-182-175A-57
; Sequence 57, Application US/08182175A
; Patent No. 5559223
; GENERAL INFORMATION:
; APPLICANT: Saverio Carl Falco
; APPLICANT: Sharon J. Keeler
; APPLICANT: Janet A. Rice
; TITLE OF INVENTION: Synthetic Storage Proteins with Defined Structure Containing
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E.I. du Pont de Nemours and Company
; STREET: 1007 Market Street
; CITY: Wilmington
; STATE: Delaware
; COUNTRY: USA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: Macintosh System, 6.0
; SOFTWARE: Microsoft Word, 4.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/182,175A
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/743,006
; FILING DATE: 9 August 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Linda Axamethy Floyd
; REGISTRATION NUMBER: 33,692
; REFERENCE/DOCKET NUMBER: BB-1031
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (302) 992-4929
; TELEFAX: (302) 892-7949
; TELEX: 835420
; INFORMATION FOR SEQ ID NO: 57:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 77 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-182-175A-57

Query Match 10.7%; Score 59.5; DB 1; Length 77;
Best Local Similarity 25.3%; Pred. No. 26;
Matches 22; Conservative 15; Mismatches 27; Indels 23; Gaps 3;

